

A Framework for Non-Asymptotic Quantum Information Theory

A dissertation submitted to

ETH ZURICH

for the degree of
Doctor of Sciences

presented by

Marco Tomamichel
Dipl. Ing. ETH

born March 13, 1981, in St. Gallen
citizen of Bosco Gurin, TI, Switzerland

accepted on the recommendation of

Prof. Dr. Renato Renner, examiner
Prof. Dr. Andreas Winter, co-examiner
Prof. Dr. Amos Lapidot, co-examiner

2012

Acknowledgements

First and foremost, I want to thank my advisor, Renato Renner. I am grateful for his eagerness to supervise me and to help me with conceptual as well as technical questions. Most importantly, I appreciated his unsurmountable optimism in the face of scientific challenges. I never encountered a problem for which he did not have at least one solution idea ready.

A special thanks goes to Roger Colbeck and Frédéric Dupuis, who shared office with me as post-docs. They always generously took time to answer my questions and I learned a lot through these interactions.

I am greatly indebted to all my coauthors for giving me the opportunity to collaborate with and learn from them. This includes Mario Berta, Roger Colbeck, Frédéric Dupuis, Torsten Franz, Fabian Furrer, Nicolas Gisin, Esther Hänggi, Stefan Hengl, Charles Lim Ci Wen, Renato Renner, Christian Schaffner, Volkher Scholz, Adam Smith, Oleg Szehr, Reinhard Werner and Severin Winkler.

I am grateful to my examiners, Andreas Winter and Amos Lapidot, for their useful comments that helped improve the final version of this thesis. Andreas Winter's suggestions led to a significant improvement of the asymptotic equipartition property. Matthias Christandl proposed to consider chained uncertainty relations. Frédéric Dupuis and Johan Åberg were bothered by me with various technical questions during the writeup, and always had an answer. Mario Berta, Dejan Dukaric, Philippe Faist, Lúcia del Rio, Cyril Stark, and Severin Winkler all gave me helpful comments on an earlier version of the thesis.

Finally, I want to thank all members of the Quantum Information Science group, past and present, for making my time at the Institute for Theoretical Physics an enjoyable one.

Abstract

This thesis consolidates, improves and extends the smooth entropy framework for non-asymptotic information theory and cryptography.

We investigate the conditional min- and max-entropy for quantum states, generalizations of classical Rényi entropies. We introduce the purified distance, a novel metric for unnormalized quantum states and use it to define smooth entropies as optimizations of the min- and max-entropies over a ball of close states. We explore various properties of these entropies, including data-processing inequalities, chain rules and their classical limits. The most important property is an entropic formulation of the asymptotic equipartition property, which implies that the smooth entropies converge to the von Neumann entropy in the limit of many independent copies. The smooth entropies also satisfy duality and entropic uncertainty relations that provide limits on the power of two different observers to predict the outcome of a measurement on a quantum system.

Finally, we discuss three example applications of the smooth entropy framework. We show a strong converse statement for source coding with quantum side information, characterize randomness extraction against quantum side information and prove information theoretic security of quantum key distribution using an intuitive argument based on the entropic uncertainty relation.

Contents

	Page
Contents	1
Abbreviations and Notation	5
1 Introduction	7
1.1 Quantum Mechanics	7
1.2 Information Theory	9
1.3 Quantum Information Theory	12
1.4 Non-Asymptotic Information Theory	14
1.5 Goal and Outline	18
2 Preliminaries	21
2.1 Linear Algebra on Hilbert Spaces	21
2.1.1 Hilbert Spaces	21
2.1.2 Operators on Hilbert Spaces	24
2.1.3 Positive Semi-Definite Operators	29
2.1.4 Operators on Tensor Spaces	30
2.1.5 Completely Positive Maps	31
2.2 Quantum Mechanics	34
2.2.1 Systems and States	35
2.2.2 Separated Systems	36
2.2.3 Evolutions and Measurements	39
2.3 Mathematical Toolkit	41
2.3.1 Semi-Definite Programs	41
2.3.2 Operator Monotone Functions	42
3 The Purified Distance	45
3.1 Introduction and Related Work	45
3.1.1 Main Contributions	46
3.1.2 Outline	46
3.2 Two Metrics for Quantum States	47
3.2.1 Generalized Trace Distance	47

CONTENTS

3.2.2	Generalized Fidelity	47
3.2.3	Purified Distance	49
3.3	Properties of the Purified Distance	50
3.4	Notational Conventions	53
4	Min- and Max-Entropies	55
4.1	Introduction and Related Work	55
4.1.1	Main Results	57
4.1.2	Outline	58
4.2	Min- and Max-Entropies	58
4.2.1	The Min-Entropy	58
4.2.2	The Max-Entropy	60
4.2.3	The von Neumann Entropy	62
4.3	Relation to Classical Entropies	63
4.3.1	Classical Min-Entropy	63
4.3.2	Classical Max-Entropy	63
4.4	Guessing Probability	64
4.5	Properties of the Min- and Max-Entropy	65
4.5.1	First Bounds and Order	66
4.5.2	Continuity	68
4.5.3	Conditioning on Classical Information	69
4.5.4	Concavity of the Max-Entropy	70
5	Smooth Entropies	71
5.1	Introduction and Related Work	71
5.1.1	Main Contributions	72
5.1.2	Outline	72
5.2	Smooth Min- and Max-Entropies	72
5.2.1	The Smooth Entropies	73
5.2.2	Remarks on Smoothing	74
5.3	Properties of the Smooth Entropies	76
5.3.1	Invariance under Isometries	76
5.3.2	Duality of Smooth Entropies	77
5.3.3	Relation between Smooth Entropies	78
5.4	Data Processing Inequalities	79
5.5	Classical Information	80
5.5.1	Smoothing of Classical States	81
5.5.2	Entropy of Classical Information	82
5.5.3	Conditioning on Classical Information	83
5.5.4	Functions on Classical Registers	84
5.6	Chain Rules	85

6	The Quantum Asymptotic Equipartition Property	87
6.1	Introduction and Related Work	87
6.1.1	Main Contributions	90
6.1.2	Outline	91
6.2	Sketch of a Classical Proof	92
6.3	Quantum Relative Entropies	93
6.3.1	Quasi-Entropies	93
6.3.2	Relative Rényi entropies	93
6.3.3	Relative Min- and Max-Entropy	94
6.4	Lower Bounds on the Smooth Min-Entropy	94
6.5	From Rényi to von Neumann Entropies	97
6.6	The Asymptotic Equipartition Property	100
6.6.1	Direct Part	100
6.6.2	Converse Part	101
7	Uncertainty Relations for Smooth Entropies	103
7.1	Introduction and Related Work	103
7.1.1	Main Contributions	105
7.1.2	Outline	108
7.2	Traditional Formulation	109
7.2.1	Overlap	109
7.2.2	Uncertainty Relation for Min- and Max-Entropies	110
7.2.3	Modeling the Measurement Basis Explicitly	111
7.3	The Generalized Uncertainty Relation	112
7.3.1	Effective Overlap	112
7.3.2	The Generalized Uncertainty Relation	113
7.4	Miscellaneous Uncertainty Relations	118
7.4.1	Commuting Measurements	118
7.4.2	Modeling the Measurement Basis Explicitly	119
7.4.3	The von Neumann Limit	120
7.4.4	The Quantum Key Distribution Setup	121
7.5	Bipartite Uncertainty Relations	122
7.5.1	Chained Uncertainty Relations	124
8	Applications	127
8.1	Full Characterization of Source Compression	127
8.1.1	One-Shot Characterization	128
8.1.2	Finite Block Lengths and Asymptotic Limits	129
8.1.3	Strong Converse	131
8.2	Randomness Extraction	132
8.2.1	Direct Bound	133
8.2.2	Converse Bound	133
8.3	Security in Quantum Key Distribution	134
8.3.1	The Protocol	134

CONTENTS

8.3.2 Security in the Finite Key Regime	135
9 Conclusions and Outlook	139
9.1 Applications of the Framework	139
9.2 Outlook and Open Questions	140
Bibliography	141
A Various Lemmas	153
A.1 Two Lemmas for Tensor Spaces	153
A.2 Entropies of Coherent Classical States	154
A.3 Selected Relations between Entropies	155
B Properties of Quasi-Entropies	159
B.1 Properties of Quasi-Entropies	159
B.2 Properties of the Rényi Entropy	160
Index	163

Abbreviations and Notation

Abbreviation	Description
CPM	Completely positive map
TP-CPM	Trace-preserving completely positive map
SDP	Semi-definite program
POVM	Positive operator-valued measurement
QKD	Quantum key distribution
i.i.d.	Independent and identically distributed
UCR	Uncertainty relation
CQ	Classical-quantum (e.g. CQ-states)
lhs.	Left-hand side (of an equation)
rhs.	Right-hand side (of an equation)

Table 1: List of Abbreviations.

Symbol	Description
A, B', AD	Typical physical systems and joint systems
X, \hat{Y}, Z	Typical registers (random variables)
$\mathcal{H}_{AB}, \mathcal{H}_X$	Hilbert spaces corresponding to a joint quantum and to a classical system
d_A	Dimension of the system A , $d_A = \dim \{\mathcal{H}_A\}$
$\text{tr}_A \equiv \text{tr}_{\mathcal{H}_A}$	Partial trace over subsystem A
$\mathcal{S}_=(\mathcal{H})$	Normalized quantum states on \mathcal{H}
$\mathcal{S}_\leq(\mathcal{H})$	Sub-normalized states on \mathcal{H}
$\rho, \hat{\tau}, \omega_A^i, \bar{\sigma}_{BC}$	Typical (possibly sub-normalized) quantum states
$ \psi\rangle, \varphi_{ABC}$	Typical pure states, i.e. rank 1 quantum states
$\pi_A, \gamma_{AA'}$	Completely mixed state on A and maximally entangled state between A and A'
$D(\rho, \tau)$	Generalized trace distance between ρ and τ
$F(\rho, \tau)$	Generalized fidelity between ρ and τ
$P(\rho, \tau)$	Purified distance between ρ and τ

Table 2: Notational Conventions for Quantum Mechanics.

CONTENTS

Symbol	Description
$\log \equiv \log_2$	Binary logarithm, i.e. logarithm to the basis 2
\ln	Natural logarithm
\mathbb{R}, \mathbb{C}	Real and complex numbers
$\mathcal{H}, \mathcal{H}'$	Typical Hilbert spaces
$ \varphi\rangle, \nu\rangle, \vartheta_i\rangle$	Typical elements of a Hilbert space, kets
$\langle\psi , \langle\phi $	Typical functionals on a Hilbert space, bras
$\text{tr}, \text{tr}_{\mathcal{H}}$	Trace and partial trace over \mathcal{H}
$\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}, \mathcal{H}')$	Linear operators on \mathcal{H} and from \mathcal{H} to \mathcal{H}'
$\mathcal{H}(\mathcal{H}), \mathcal{P}(\mathcal{H})$	Hermitian and positive semi-definite operators on \mathcal{H}
$A \geq B$	Equivalent to $A - B \in \mathcal{P}(\mathcal{H})$
X^\dagger	The adjoint operator of X
X^T	The transpose operator of X , defined with regards to a basis that needs to be specified
X^{-1}	The generalized inverse operator of X
$\ X\ _\infty$	Maximum singular value of X , operator norm
$\mathbb{1}$	Identity operator
E, X, L	Typical linear operators on a Hilbert space
M, N	Typical positive semi-definite operators
U, V, W	Typical unitary operators or isometries
P, Π	Typical projectors
$\mathcal{E}, \mathcal{F}, \mathcal{I}$	Typical trace-preserving completely positive maps

Table 3: Notational Conventions for Mathematical Expressions.

Chapter 1

Introduction

This chapter starts with a rather philosophical introduction into quantum mechanics that does not assume any prior knowledge. Here, we attempt to explain and justify the information theoretic approach this thesis will take on the topic. This will lead into a short overview over some relevant aspects of information and quantum information theory. Then, we discuss the importance of non-asymptotic quantum information theory to characterize elementary information processing tasks. Finally, the introduction ends with a detailed outline of the thesis.

1.1 Quantum Mechanics

The laws of *quantum mechanics* govern the behavior of microscopic physical systems and are verified daily in experiments conducted in physics laboratories worldwide. Here, we restrict our attention to non-relativistic quantum mechanics and take a static perspective on quantum theory, where the state of a physical system is the central object of interest.

Non-relativistic quantum mechanics is the theoretical basis of today's semiconductor industry. As we further miniaturize physical devices used for information processing, the impact of quantum mechanics will become more and more relevant. Hence, a thorough understanding of quantum physics will be pivotal to successfully engineer the next generation of information processing devices. In quantum cryptography [BB84, Eke91], some of the more peculiar effects of quantum mechanics are already exploited today in order to ensure secrecy of the communication between two distant parties.

While an understanding of the laws of quantum mechanics is thus necessary in order to fully comprehend the physical world surrounding us, these laws are nevertheless in stark contrast with our intuition about the causal structure of the universe. In our everyday experience, any observation about a physical system can be predicted perfectly given a complete

1. INTRODUCTION

objective description of the state of said system. For example, position, momentum and spin of a football are part of its objective state and every *observer* can verify their value independently, given appropriate measurement equipment. This perspective is deeply ingrained in our language: We talk about the position, momentum and spin of a football, directly associating our observations (the outcomes of position, momentum and spin measurements) with the football and, thus, implying that there exists an objective reality — the state of the football — beyond our observations.

However, quantum mechanics does not allow for an objective description of the state of a physical system that deterministically predicts all observations about said system. For example, it is impossible to write down a quantum mechanical state of an electron such that the outcomes of both position and momentum measurements can be predicted with arbitrary precision. This is known as the Heisenberg *uncertainty principle* [Hei27] and has mystified physicists since the early days of quantum mechanics. In what sense can position and momentum then be considered real? Is it even permissible to speak of the position and momentum of a quantum mechanical object like an electron?

Furthermore — as if the loss of determinism was not enough to confuse our human minds — quantum mechanics generally does not even allow for an objective description of the state of a physical system that determines the *probabilities* with which different observations about it are made. These probabilities, even if the description of the state is complete within quantum mechanics, are in general subjective to the observer. It is crucial to note here that classical theory allows to describe subjective knowledge about the state of a system, for example through conditional probability distributions. (We will encounter an example model of subjective classical information in the next section.) However, such a description cannot be considered complete within classical theory, since, in principle, every observer may hold a copy of the full objective state of the system, e.g. position, momentum and spin of the football.

To convince yourself that quantum mechanics is incompatible with this notion of objectivity, consider two particles that are in an *entangled* state. More specifically, this could be two electrons, A and B , that are in a spin singlet state. In this setting, quantum mechanics postulates that an observer controlling electron B can perfectly predict the outcomes of all possible spin measurements on electron A .¹ We call this observer an *omniscient* observer of the electron A . At the same time, all other observers are *ignorant* about A , which means that they will see all possible measurement outcomes with equal probability. Quantum mechanics, in contrast to classical theory, does not allow to copy the quantum information encoded in

¹The observer simply measures the spin of B in the same direction as the spin of A is measured to get a fully anti-correlated result.

the state of electron B and share it with other observers. On an intuitive level, this is often explained using the concepts of *no-cloning* [WZ82] or the *monogamy of entanglement*. These statements can be formulated quantitatively and we will see in Chapter 4 that the fact that one observer is close to omniscient about a quantum system implies that all other observers are necessarily close to ignorant about it.

This apparent deficiency of quantum mechanics famously led Einstein, Podolsky and Rosen to ask “Can quantum-mechanical description of physical reality be considered complete?” [EPR35]. The search for a theory, consistent with quantum mechanics, that would assign objective descriptions¹ to physical systems that probabilistically predict observations about said systems, was abandoned when Bell [Bel64] as well as Kochen and Specker [KS68] proved their pivotal theorems. They show that such objective descriptions cannot fully predict the probabilities of certain observations that are consistent with quantum mechanics and can be verified experimentally.²

Furthermore, the existence of preferred observers renders *quantum cryptography* possible. In quantum key distribution, we consider two players, traditionally called Alice and Bob, who want to share a secret over a public channel. Once the two players can establish that Bob is almost *omniscient* about a quantum system Alice controls, a shared secret can be produced by an arbitrary measurement of Alice’s system: while Bob can predict Alice’s observed measurement result well, by the laws of quantum mechanics, any eavesdropper is guaranteed to have almost no knowledge about it.

1.2 Information Theory

Luckily, some of the counter-intuitive effects of quantum mechanics—in particular, the appearance of non-determinism and subjective knowledge as explained above—have also been considered in a completely different context, in probability and *information theory*. The latter was founded by Nyquist [Nyq28] and Shannon [Sha48] in the early days of telegraph communication in order to investigate information sources and the capacities of channels to transmit digital information.

Following Shannon, consider a source that outputs English text. Clearly, the different letters in the Latin alphabet will not be produced with equal probability by such a source. This is used, for example, in the Morse

¹Such objective descriptions of a system are usually called (*local*) *hidden variables* in this context, since they are not accessible from within quantum mechanics.

²In fact, it was recently shown that if such objective descriptions have any predictive power exceeding quantum mechanics, they can be falsified experimentally under minimal assumptions. [CR08, CR10]

1. INTRODUCTION

alphabet, where the most frequent letters are given the simplest codes.¹ Going one step further, note that certain combinations of letters are more likely than others. For example, the letter ‘q’ is almost always followed by a ‘u’. Hence, given that the last letter this source produced was a ‘q’, the information content, or *uncertainty*, of the next letter is very low.

This implies that the probability with which letters occur is relevant in order to characterize the information content of the source — a source that produces every letter with equal probability produces more information than a source that is biased towards certain letters. To quantify these considerations, Shannon used *entropies*. Very generally, he considered *events* that are known to occur with a certain probability p and assigned to them the value $-\log p$, the *surprisal*, which measures how surprising the occurrence of the event is. (The logarithm is taken to the binary basis throughout this thesis.) If an event occurs with certainty its surprisal is zero and as the event gets less likely its surprisal can grow arbitrarily large. In the above example, the surprisal of the letters ‘e’ and ‘t’ is smaller than the surprisal of less common letters. Moreover, the surprisal of the letter ‘u’ given that the last letter was ‘q’ is close to zero.

The *Shannon entropy* of a source is the average surprisal of the events it produces and is introduced more formally in the following. For the purpose of this thesis, a *probability distribution* on a *random variable* X that takes values from a discrete set \mathcal{X} is a function P_X from \mathcal{X} to non-negative real numbers with the property that $\sum_{x \in \mathcal{X}} P_X(x) = 1$. The Shannon entropy of a random variable X with distribution P_X is then defined as

$$H(X)_P := - \sum_{x \in \mathcal{X}} P_X(x) \log P_X(x). \quad (1.1)$$

Here, $-\log P_X(x)$ is the surprisal of the event “ $X = x$ ” and has the properties described above. On one hand, Shannon [Sha48] and later Rényi [Rén61] derived the mathematical form of the entropy, Eq. (1.1), from intuitive axioms that measures of the average surprisal should satisfy. On the other hand, the Shannon entropy is intimately related to physics and it was inspired directly by the Gibbs entropy of thermodynamics [Gib76, Bol72].

Next, we consider the case where we have *side information* about X , modeled as another random variable Y . The two random variables, X and Y , have a joint probability distribution, $P_{XY}(x, y)$. The *marginal* probability distributions of the individual random variables X and Y are thus given as $P_X(x) = \sum_y P_{XY}(x, y)$ and $P_Y(y) = \sum_x P_{XY}(x, y)$, respectively. We employ the *conditional Shannon entropy* of X conditioned on Y , which is defined as $H(X|Y)_P := H(XY)_P - H(Y)_P$. This definition, aside from

¹The most frequent letters ‘e’ and ‘t’ are given the codes ‘.’ and ‘—’, respectively, whereas all other letters have codes with at least two symbols.

its operational interpretation that we will discuss below, has very natural properties that we would expect from a conditional entropy, i.e. a measure about the average *surprisal* or uncertainty of X given *side information* Y . For example, we have $H(X|Y)_P \leq H(X)_P$, namely, the uncertainty about X increases if one ignores the side information. Equality holds if the side information is independent of X , i.e. if $P_{XY}(x, y) = P_X(x)P_Y(y)$.¹ Relations of this type are called *data processing* inequalities and will be discussed in Chapter 5.

Conditional entropies can be employed to model subjective classical information.² Consider the example of an unbiased die that is thrown secretly in a cup. The state of the die before the result is revealed is modeled as a random variable, X , on the set $\{\square, \square, \square, \square, \square, \square\}$ of faces with probability distribution $P_X(\cdot) = \frac{1}{6}$. The Shannon entropy or uncertainty about the outcome is simply $H(X)_P = \log 6 \approx 2.58$. However, an *observer* may own a futuristic device, Y , that correctly predicts whether the value is even or odd with probability $\frac{3}{4}$ from the sound the die emits when thrown. The correlations between the result of the throw and this device can be described by a joint probability distribution P_{XY} satisfying, e.g., $P_{XY}(\square, \text{"odd"}) = \frac{1}{6} \cdot \frac{3}{4} = \frac{1}{8}$ and $P_{XY}(\square, \text{"even"}) = \frac{1}{6} \cdot \frac{1}{4} = \frac{1}{24}$. The uncertainty this observer has about the outcome is reduced to $H(X|Y)_P \approx 2.39$. Another observer may own an even more powerful device, Z , that predicts the value of X with certainty; hence, $H(X|Z)_P = 0$. The subjective uncertainty an observer has about the outcome of the throw thus critically depends on the available *side information*, Y or Z .

The Shannon entropy has found a vast number of applications in information theory. Shannon's source coding theorem [Sha48] concerns itself with the question of how much we can compress the output of a source. If a source produces a long stream of *independent and identically-distributed* (i.i.d.) symbols X , it states that any attempt to compress the output to less than $H(X)_P$ logical bits per symbol will almost certainly lead to information loss. Here, P_X is the probability distribution of the source symbols. On the other hand, it is possible to compress the stream to arbitrarily close to $H(X)_P$ bits per symbol with insignificant probability of informa-

¹More generally, and in accordance with our intuition, any function applied to the random variable Y can at most increase the uncertainty about X , i.e. $H(X|Y)_P \leq H(X|Z)_P$ if $Z = f(Y)$. The entropy $H(X|Z)_P$ is evaluated for the probability distribution induced by f , i.e. the distribution $P_{XZ}(x, z) = \sum_{y: f(y)=z} P_{XY}(x, y)$.

²For another example, we return to Shannon's text sources and treat them quantitatively. Let us first consider a source that outputs all 26 letters of the English alphabet with equal probability. Clearly, the entropy of the output U of this source is $H(U) = \log 26 \approx 4.7$. A source that outputs letters of English text, X , has lower entropy — $H(X) \approx 4.14$ [Sha51] — due to the non-uniform distribution of the different letters. Moreover, if we consider two consecutive letters of English text, denoted X_1 and X_2 , we find that the conditional entropy — $H(X_2|X_1) \approx 3.56$ [Sha51] — is even smaller due to correlations between the probabilities of adjacent letters.

1. INTRODUCTION

tion loss. More generally, if some additional side information Y about each X is available, the quantity $H(X)_P$ can be replaced by $H(X|Y)_P$ in the above statements [SW73], where P_{XY} is the joint probability distribution of source and side information. The Shannon entropy thus quantifies the amount of memory (in bits) needed to store the output of a source reliably. Memory can be considered a *resource* in source coding and, more generally, the Shannon entropy is often used to analyze the resource usage of a task in the limit of many independent and identical repetitions. In the following, we call this the *i.i.d. limit*.

As a further important example, the capacity of a channel to transmit information (in the i.i.d. limit of many independent uses of a memoryless channel) can be expressed in terms of Shannon entropies [Sha48].

1.3 Quantum Information Theory

An observation about a physical system, in the sense described above, constitutes an *event*. Moreover, a *measurement*, i.e. a complete set of mutually exclusive observations about a physical system, can be seen as an information source. It is thus very natural to try to model quantum measurements using the entropy formalism of the previous section. Moreover, many information theoretic tasks (we will discuss the example of source compression in detail) can be generalized to the quantum setting and thus the question arises whether the resource usage in these tasks can be characterized using an analogue of the Shannon entropy as well. We will see that this is indeed possible.

In the quantum formalism, to be formally introduced in Chapter 2, a random variable is modeled as a (classical) *register*, X and the probability distribution, P_X , is represented as a diagonal matrix with the probabilities $P_X(x)$ as entries. This is a special case of a *quantum system*. More generally, the state of a quantum system, A , is modeled as a positive semi-definite operator with unit trace, ρ_A , called the density operator or *quantum state*. Schumacher and Wootters [Sch95] introduced the term *qubit* to denote the smallest unit of quantum information. The spin degree of freedom of an electron, as we have seen before, constitutes a physical example of such a qubit system. Holevo's bound [Hol73] implies that a qubit can only store one bit of classical information. This indicates that the power of quantum information over classical information lies in the possible correlations between multiple quantum systems and not necessarily in the quantum systems themselves.

The *von Neumann entropy* [vN32] of a quantum system is given as

$$H(A)_\rho := -\log \operatorname{tr}(\rho_A \log \rho_A),$$

where tr denotes the trace. For a register X this expression reduces to

1.3 Quantum Information Theory

the *Shannon entropy*. A measurement, in this formalism, is simply a map from a quantum system to a register.

To see a first application of the von Neumann entropy in quantum information theory, we consider a simple *uncertainty relation*. For example, let A be qubit, for example an electron spin in a state ρ_A and let X be a register containing the outcome of either one of two *incompatible measurements* that are chosen at random, for example spin measurements in the x or y direction. Then, an entropic formulation of Heisenberg's uncertainty principle tells us that $H(X)_\rho \geq \frac{1}{2}$. [MU88]. Hence, independently of how the initial state ρ_A is prepared (as long as there is no quantum side information present), there will be uncertainty about at least one of the outcomes. Such uncertainty relations, expressed in terms of entropies, are the topic of Chapter 7.

In order to investigate the effect of quantum correlations, we introduce the conditional von Neumann entropy. For a bipartite quantum state ρ_{AB} , this is given as $H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$. If the system A is a classical register, this entropy is nonnegative and can safely be interpreted as a measure of uncertainty given quantum side information. More explicitly, we reconsider the example where two electrons, A and B , are in a spin singlet state. As we have noted before, quantum mechanics predicts that a preferred observer controlling the electron B can predict any measurement outcome on the A electron with certainty. Thus, the *surprisal* of any observation about electron A is zero for this observer. More formally, we denote by X the random variable that stores the outcome of an arbitrary spin measurement on the electron A . Then, the conditional *von Neumann entropy* evaluates to $H(X|B)_\rho = 0$. Here, the entropy is evaluated for the post-measurement state ρ_{XB} that results from measuring X on the joint *quantum state* ρ_{AB} . On the other hand, all other observers will see a uniformly random measurement outcome and thus have maximum surprisal. We denote by C such an *ignorant* observer. And indeed, for any tripartite quantum state ρ_{ABC} that is compatible with the marginal state ρ_{AB} , we find $H(X|C)_\rho = 1$. Again, the entropy is evaluated for the post-measurement state ρ_{XC} that results from measuring X on the joint state ρ_{AC} .

However, the conditional von Neumann entropy can also be evaluated for the quantum state ρ_{AB} of the electron pair before measurement. This entropy evaluates to $H(A|B)_\rho = -1$ in the example above. Generally, the conditional entropy can be negative in the presence of *entanglement* and its interpretation as a measure of uncertainty has been controversial. (See, e.g., [HOW05, BCC⁺10, dRAR⁺11] for recent work elucidating the issue.)

Generalizing the source coding theorem with side information of the previous section, we ask how much we can compress a long stream of *i.i.d.* random variables X such that an observer with quantum side information B about each X can reconstruct the original from the compressed stream.

1. INTRODUCTION

Such side information can be modeled in a joint quantum state ρ_{XB} . The answer was given by Devetak and Winter [DW03], who propose a quantum generalization of the Slepian-Wolf theorem. They show that, consistent with the classical result, the stream can be compressed to $H(X|B)_\rho$ bits per symbol in the *i.i.d. limit* of long streams.

In another generalization of source coding, we ask how much we can compress a long stream of independent and identical quantum systems A that are in the state ρ_A . The answer, that such a string can be compressed to $H(A)_\rho$ qubits, was given by Schumacher [Sch95] in his pioneering work on quantum information theory.

We can go one step further into the quantum world and consider *state merging* [HOW06]. Here, we start with joint quantum state ρ_{AB} shared between two parties, A and B . The task is to recreate the state (including its correlations with the environment) at B using free classical communication between A and B and *entanglement* between A and B as a *resource*. The amount of this resource that needs to be utilized (in the *i.i.d. limit* of many repetitions of the task) is then quantified by the conditional von Neumann entropy, $H(A|B)_\rho$. The conditional entropy can be consistently interpreted as the amount of entanglement needed to complete the task. In particular, if $H(A|B)_\rho$ is negative (which may happen only in the presence of entanglement in ρ_{AB}), it is possible to extract $-H(A|B)_\rho$ units of entanglement from the state while recreating it at B .

1.4 Non-Asymptotic Information Theory

So far, we have considered tasks in the *i.i.d. limit* of many repetitions and found that the resource usage is characterized by expressions involving the Shannon and von Neumann entropies. This leads to the question of what happens to the resource usage when we consider a finite number of trials and allow a small probability of failure. In many context, in particular in cryptography, we are even interested in the amount of resource needed to perform a task just once, with high probability of success. We call this the *one-shot* setting in the following. These questions, which we are going to tackle in the quantum setting, have also been the topic of recent research in classical information theory (see, e.g., [PPV10, Pol10, WCR09]).

It turns out that the Shannon entropy and von Neumann entropies, which have been invaluable tools in the previous sections, are insufficient to characterize the required *resources* in the one-shot setting. Nevertheless, there are other entropies that can be used instead to quantify uncertainty, as we will see in the following.

We will now introduce some aspects of non-asymptotic information theory on the example of *source compression*. For this purpose, let us consider a source X with probability distribution P_X . The number of bits

1.4 Non-Asymptotic Information Theory

of memory needed to store the output of this source so that it can be recovered with certainty is given by $\lceil H_0(X)_P \rceil$, where $H_0(X)_P$ denotes the *Hartley entropy* [Har28] of the distribution P_X , defined as

$$H_0(X)_P := \log |\{x : P_X(x) > 0\}|,$$

The Hartely entropy corresponds to the *Rényi entropy* of order 0 [Rén61] and simply measures the size of the support of X . As an example, we consider again a source that outputs characters of the English alphabet. If we want to store a single character produced by this source such that it can be recovered with certainty, we clearly need $\lceil \log 26 \rceil = 5$ bits of memory as a *resource*. This first result is rather unsatisfactory since the resource usage does not depend on the actual probability distribution of the letters but only on the size of the alphabet.

More interestingly, we may ask how much memory we need to store the output if we allow a small probability of failure, ε . One way to tackle such problems is by investigating encoders that assign code words (i.e. binary strings) of a fixed length m (in bits) to the events a source produces. These code words are then stored in m bits of memory and a decoder is later used to compute an estimate of X from that memory. For a source X with probability distribution P_X , we are thus interested in the minimum code length, $m^\varepsilon(X)_P$, for which there exists an encoder and a decoder that achieve a probability of failure not exceeding ε . Gallager [Gal79] showed that a random assignment of source events to code words on average leads to a failure probability of at most ε if the code length is sufficiently long. His results imply that the minimal code length satisfies

$$m^\varepsilon(X)_P \leq H_{\max}(X)_P + \log \frac{1}{\varepsilon} + 1. \quad (1.2)$$

Here, $H_{\max}(X)_P$ denotes the *max-entropy*, which corresponds to the *Rényi entropy* of order $\frac{1}{2}$ and is defined as

$$H_{\max}(X)_P := \log \left(\sum_{x \in \mathcal{X}} \sqrt{P_X(x)} \right)^2.$$

Upper bounds of the type (1.2) are called *direct bounds* and show that there exist protocols using a certain amount of resource that do not exceed a fixed probability of failure.

This analysis can be further refined by *smoothing* the max-entropy. For this purpose, let us consider probability distributions Q_X that are close to P_X and have max-entropy $H_{\max}(X)_Q$ smaller than $H_{\max}(X)_P$. In principle, we could design the encoding and decoding scheme for a source with distribution Q_X instead of P_X . Clearly, this reduces the upper bound in (1.2). In fact, the upper bound can be expressed in terms of a *smooth max-entropy*,

1. INTRODUCTION

$H_{\max}^{\varepsilon_1}(X)_P := \inf_{Q \approx P} H_{\max}(X)_Q$, which minimizes the max-entropy over probability distributions Q_X that are within *statistical distance* ε_1 of P_X .¹ More precisely, they satisfy $D(P_X, Q_X) := \frac{1}{2} \sum_x |P_X(x) - Q_X(x)| \leq \varepsilon_1$.

However, if this scheme is applied to the original source, we incur an additional error that depends on the statistical distance between P_X and Q_X . The total probability of failure, $\varepsilon \leq \varepsilon_1 + \varepsilon_2$, is thus split into two contributions: the statistical distance of the distributions, $D(P_X, Q_X) \leq \varepsilon_1$, and the contribution from the Gallager bound, which we denote ε_2 . Hence, we get improved direct bounds on the minimal code length:

$$m^\varepsilon(X)_P \leq H_{\max}^{\varepsilon_1}(X)_P + \log \frac{1}{\varepsilon_2} + 1, \quad \forall \varepsilon_1, \varepsilon_2 \text{ with } \varepsilon_1 + \varepsilon_2 = \varepsilon. \quad (1.3)$$

We also consider *converse bounds* that give a lower bound on the resources required to achieve a certain probability of success. In fact, it can be shown that

$$m^\varepsilon(X)_P \geq H_{\max}^\varepsilon(X)_P. \quad (1.4)$$

Hence, both the lower and upper bound on the quantity $m^\varepsilon(X)_P$ can be expressed in terms of a smooth max-entropy. We thus say that the required memory for *one-shot* source compression is *characterized* by the smooth max-entropy. (Note that a more detailed analysis of source compression with quantum side information can be found in Chapter 8 and [RR10].)

To see why the Shannon entropy does not suffice to characterize the *one-shot* version of source compression, consider a source that produces the symbol ‘a’ with probability $\frac{1}{2}$ and k other symbols each with probability $\frac{1}{2k}$. On one hand, for any fixed failure probability, the converse bound in (1.4) evaluates to approximately $\log k$ for large enough k . This implies that we cannot compress this source much beyond the *Hartley entropy*. On the other hand, the Shannon entropy of this distribution is $\frac{1}{2}(\log k + 2)$ and underestimates the required resources by a factor two.

In this thesis, we will mainly encounter two entropic quantities that are defined for quantum states in Chapters 4 and 5. Surprisingly, it turns out that these two entropies suffice to characterize the resource usage of many information theoretic tasks in the *one-shot* setting in a manner similar to source compression. The first quantity, the *smooth min-entropy*, H_{\min}^ε , is a generalization of the min-entropy or Rényi entropy of order ∞ , which evaluates the minimum *surprisal* of a random variable X . Namely,

$$H_{\min}(X)_P := \min_{x \in \mathcal{X}} -\log P_X(x),$$

¹Note that we use the statistical distance as a metric here for convenience of exposition. In Chapter 3, we will argue that the *purified distance* should be used instead to define the smooth entropies in the quantum setting. All smooth entropies used in this thesis, except in this introductory exposition, are thus based on the purified distance as a metric.

One of the major applications of the smooth min-entropy is in *randomness extraction*, where it characterizes the amount of uniform randomness that can be extracted from a biased source [ILL89]. The second quantity, the *smooth max-entropy*, H_{\max}^ε , is used to characterize source compression, as we have seen above.

The smooth min- and max-entropies were first generalized to the quantum settings by Renner and König [RK05, Ren05]. The smooth min-entropy we use in this thesis is a refined version of the smooth min-entropy proposed in [Ren05], whereas the smooth max-entropy is based on later work by Renner, König and Schaffner [KRS09]. The significance of these extensions to the quantum setting stems from their operational meaning. For example, the quantum generalization of the min-entropy, $H_{\min}^\varepsilon(X|B)_\rho$, characterizes randomness extraction against quantum side information B , i.e. it characterizes the amount of uniform randomness, independent of the side information B , that can be extracted from X [Ren05, TSSR11]. The quantum generalization of the max-entropy, $H_{\max}^\varepsilon(X|B)_\rho$, characterizes source compression with quantum side information B in the *one-shot* setting [RR10]. Moreover, the amount of entanglement needed in one-shot *state merging* is characterized by $H_{\max}^\varepsilon(A|B)_\rho$ [Ber08]. We will discuss source compression with quantum side information as well as randomness extraction against quantum side information in Chapter 8.

In addition to their operational meaning, the smooth entropies exhibit many useful properties, including *data processing* inequalities and *chain rules*. Furthermore, the smooth min- and max-entropies converge to the von Neumann entropy in the *i.i.d. limit*. For any $0 < \varepsilon < 1$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^\varepsilon(X^n|B^n)_\rho = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\max}^\varepsilon(X^n|B^n)_\rho = H(X|B).$$

We call this the entropic form of the *asymptotic equipartition property* and it is the topic of Chapter 6.

This means that if the resource usage is characterized by a smooth entropy in the *one-shot* setting, the resource usage in the *i.i.d. limit* is given by the von Neumann entropy. In fact, a simple analysis shows the upper and lower bounds in (1.3) and (1.4) on the code rate, $r = \frac{1}{n} \log m^\varepsilon(X^n)$, converge to the von Neumann limit when we let n go to infinity for any fixed $0 < \varepsilon < 1$. This also shows the advantage of the smoothed direct bound (1.3) over the Gallager bound (1.2), for which such an asymptotic convergence can only be shown using additional techniques.

The smooth entropies satisfy a *duality* relation. For any tripartite quantum state ρ_{ABC} and any $0 \leq \varepsilon < 1$, we find [KRS09, TCR10]

$$H_{\min}^\varepsilon(A|B)_\rho \geq -H_{\max}^\varepsilon(A|C)_\rho \quad (1.5)$$

and equality holds if the joint state ρ_{ABC} is pure.¹ This relation provides

¹Pure states offer the most complete description of a joint quantum system.

1. INTRODUCTION

a connection between the min- and max-entropy, and thus the tasks characterized by them, that does not exist in classical information theory.

It also allows us to close the circle to the discussion of quantum mechanics at the beginning of this chapter. For this purpose, let B and C be two observers of a quantum system A . Then, the min-entropy, $H_{\min}(A|B)$, can be viewed (cf. Chapter 4) as the distance of B to an *omniscient observer* of the quantum system A . Furthermore, the max-entropy, $-H_{\max}(A|C)$, can be viewed as the distance of C to an *ignorant observer* of A . The duality relation, Eq. (1.5), thus states that if B is close to an omniscient observer of the quantum system A , then C is at least as close to an ignorant observer of A . This can be seen as a manifestation of the subjective knowledge of observers quantum mechanics imposes on the world.

1.5 Goal and Outline

The goal of this thesis is to consolidate the smooth entropy framework for non-asymptotic information theory and to introduce important additions to the framework, including the entropic asymptotic equipartition property and various uncertainty relations. This work should provide a reference for researchers interested in the smooth entropy framework for non-asymptotic quantum information theory. The focus of this work is thus mainly on the properties of the smooth entropies and not on their applications.

The remainder of this thesis is organized as follows.

In Chapter 2, the notation and mathematical foundations of Hilbert space quantum mechanics are introduced. Relevant results of linear algebra are summarized in Section 2.1 and the axioms of quantum mechanics are introduced in 2.2. Moreover, Section 2.3 covers operator convex functions and semi-definite programming, completing a mathematical toolkit that will be used extensively in this thesis.

In Chapter 3, we introduce a novel measure of distance between (potentially incomplete) quantum states, the purified distance. In particular, we explore its properties and argue why they are relevant for the definition of the quantum smooth entropies.

In Chapter 4, we formally define the min- and max-entropies for quantum states, give a collection of different expressions for the entropies and explore some of their properties. In particular, we explore the relation between the min- and the max-entropy and their relation to classical Rényi entropies. We also show that both entropies are continuous functions of the quantum state.

In Chapter 5, we introduce smoothing and define the smooth min- and max-entropy. Various properties of the smoothing operation are discussed. We establish relations between the smooth min- and max-entropy and investigate data processing inequalities. The special case where one or

more systems are classical is considered in detail. We also give a list of chain rules that have recently been shown for the smooth min- and max-entropies.

In Chapter 6, we show that the smooth entropies converge to the von Neumann entropy when we consider a sequence of independent and identically distributed quantum systems. This result is an entropic version of the asymptotic equipartition property and confirms the fundamental role of the von Neumann entropy in quantum information theory. In addition to the asymptotic result, the chapter also provides bounds for finite block lengths.

In Chapter 7, we introduce a variety of entropic uncertainty relations that give bounds, analogous to Heisenberg's uncertainty principle, on the uncertainty of the outcomes of two incompatible measurements on a quantum system. These uncertainties are expressed in terms of smooth entropies as well as von Neumann entropies, which makes them directly applicable to problems in quantum cryptography.

In Chapter 8, we combine results from Chapters 3-6 to investigate two information theoretic tasks, source compression with quantum side information and randomness extraction against quantum side information. In particular, we show a strong converse statement for source compression which implies that any attempt to compress to less than the Shannon limit will fail with high probability. These results can then be used in another example application, quantum key distribution. There, we show how the entropic uncertainty relations of Chapter 7 can be employed in order to prove security of the original BB84 quantum key distribution protocol in a concise and intuitive way.

The thesis ends with Chapter 9 in a short conclusion and outlook. Some open problems are discussed.

1. INTRODUCTION

[TCR09]	M. Tomamichel , R. Colbeck, R. Renner. <i>The Fully Quantum Asymptotic Equipartition Property</i> . IEEE Trans. on Inf. Theory 55(12):5840–5847, 2009.
[TCR10]	M. Tomamichel , R. Colbeck, R. Renner. <i>Duality Between Smooth Min- and Max-Entropies</i> . IEEE Trans. on Inf. Theory, 56(9):4674–4681, 2010.
[TSSR11]	M. Tomamichel , C. Schaffner, A. Smith, and R. Renner. <i>Leftover Hashing Against Quantum Side Information</i> . IEEE Trans. on Inf. Theory, 57(8):5524–5535, 2011.
[TR11]	M. Tomamichel and R. Renner. <i>Uncertainty Relation for Smooth Entropies</i> . Phys. Rev. Lett., 106(11), 2011.
[WTHR11]	S. Winkler, M. Tomamichel , S. Hengl, and R. Renner. <i>Impossibility of Growing Quantum Bit Commitments</i> . Phys. Rev. Lett., 107(9), 2011.
[TLGR11]	M. Tomamichel , C. C. W. Lim, N. Gisin, and R. Renner. <i>Tight Finite-Key Analysis for Quantum Cryptography</i> . Nat. Commun. 3(634), 2012.
[HT11]	E. Hänggi and M. Tomamichel . <i>The Link between Uncertainty Relations and Non-Locality</i> . (submitted).
[SDTR11]	O. Szehr, F. Dupuis, M. Tomamichel , and R. Renner. <i>Decoupling with unitary almost two-designs</i> . (submitted).

Table 1.1: List of Publications.

Chapter 2

Preliminaries

The preliminaries consist of three sections. The first two sections, which cover the mathematical foundations of Hilbert space quantum mechanics in finite dimensions, can be skipped entirely by readers already familiar with the concepts of linear algebra used in quantum information theory. The notation used throughout this thesis is introduced in these sections and summarized in Tables 3 and 2. The last section then introduces some mathematical tools that are needed to derive the results of the following chapters, including semi-definite programming and operator monotone functions.

2.1 Linear Algebra on Hilbert Spaces

This section is based on many introductory text books, mostly on the two books *Matrix Analysis* and *Positive Definite Matrices* by Bhatia [Bha97, Bha07]. Moreover, John Watrous’s lecture notes [Wat08] as well as Nielsen and Chuang’s *Quantum Computation and Quantum Information* [NC00] were an invaluable resource.

2.1.1 Hilbert Spaces

Bras and Kets

Let \mathcal{H} be a finite-dimensional *vector space* over the complex numbers equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$. In the following, we will call \mathcal{H} a *Hilbert space*. The dual space of \mathcal{H} is the Hilbert space of (linear) functionals from \mathcal{H} to \mathbb{C} . We use Dirac’s bra-ket notation to denote elements of \mathcal{H} and its dual space, \mathcal{H}^* . Every ket, $|\psi\rangle \in \mathcal{H}$, is in one-to-one correspondence with its dual bra, $\langle\psi| \in \mathcal{H}^*$. The bra is defined in terms of the ket via the Hilbert space’s inner product as

$$\langle\psi| : |\varphi\rangle \mapsto \langle\psi|\varphi\rangle := \langle|\psi\rangle, |\varphi\rangle\rangle. \quad (2.1)$$

2. PRELIMINARIES

The right-hand side of the above equation gives a natural expression for the inner product in terms of a bra-ket product, and we will use this notation frequently. The bra-ket product $\langle \cdot | \cdot \rangle$ in (2.1) has the following properties, which follow directly from properties of the underlying inner product.

- Conjugate symmetry: $\langle \psi | \varphi \rangle = \overline{\langle \varphi | \psi \rangle}$.
- Sesquilinearity: Let $|\vartheta_1\rangle, |\vartheta_2\rangle \in \mathcal{H}$ and $\alpha_1, \alpha_2 \in \mathbb{C}$ such that $|\varphi\rangle = \alpha_1|\vartheta_1\rangle + \alpha_2|\vartheta_2\rangle$. Then, $\langle \psi | \varphi \rangle = \alpha_1\langle \psi | \vartheta_1 \rangle + \alpha_2\langle \psi | \vartheta_2 \rangle$ and, due to conjugate symmetry, $\langle \varphi | \psi \rangle = \bar{\alpha}_1\langle \vartheta_1 | \psi \rangle + \bar{\alpha}_2\langle \vartheta_2 | \psi \rangle$.
- Positive-definiteness: $\langle \psi | \psi \rangle \geq 0$ with equality if and only if $|\psi\rangle = \mathbf{0}$, where $\mathbf{0}$ is the zero element of the vector space.

Norms and Metrics

The usefulness of Hilbert spaces in physics stems in part from the fact that they have a natural measures of angle and distance through the inner product. In fact, the inner product of a Hilbert space \mathcal{H} induces a *norm*, $\|\cdot\| : \mathcal{H} \rightarrow \mathbb{R}$, which in turn induces a *metric*, $D(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$. They are given by the expressions

$$\| |\psi\rangle \| := \sqrt{\langle \psi | \psi \rangle} \quad \text{and} \quad D(|\psi\rangle, |\varphi\rangle) := \| |\psi\rangle - |\varphi\rangle \|.$$

More generally — and for later reference — a metric on an arbitrary set is defined as follows.

Definition 2.1 (Metric). *Let \mathcal{X} be a set and let $a, b, c \in \mathcal{X}$. Then, the functional $D : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is a metric on \mathcal{X} if it satisfies*

1. *Positive-definiteness: $D(a, b) \geq 0$ with equality if and only if $a = b$.*
2. *Symmetry: $D(a, b) = D(b, a)$.*
3. *Triangle inequality: $D(a, c) \leq D(a, b) + D(b, c)$.*

It is easy to verify that the induced metric introduced above fulfills these conditions.

Orthonormal Bases

Let \mathcal{B} be a set of kets $|\vartheta_1\rangle, |\vartheta_2\rangle, \dots, |\vartheta_n\rangle \in \mathcal{H}$. The *linear span* of \mathcal{B} is the subspace of \mathcal{H} containing all linear combinations these kets,

$$\text{span}\{\mathcal{B}\} = \text{span}\{|\vartheta_1\rangle, |\vartheta_2\rangle, \dots, |\vartheta_n\rangle\} := \left\{ \sum_{i=1}^n \alpha_i |\vartheta_i\rangle : \alpha_i \in \mathbb{C} \right\}.$$

2.1 Linear Algebra on Hilbert Spaces

We call \mathcal{B} a *basis* of \mathcal{H} if $\text{span}\{\mathcal{B}\} = \mathcal{H}$. Furthermore, a basis $\{|e_i\rangle\}_{i=1}^d := \{|e_1\rangle, |e_2\rangle, \dots, |e_d\rangle\}$ is called *orthonormal* if its elements are mutually orthonormal. Formally, this means that $\langle e_i | e_j \rangle = \delta_{ij}$, where the *indicator* function is given by $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Every orthonormal basis has exactly d elements, where $d := \dim\{\mathcal{H}\}$ is the Hilbert *dimension* of \mathcal{H} . A ket $|\psi\rangle \in \mathcal{H}$ has a unique *decomposition* into any orthonormal basis $\{|e_i\rangle\}$, where it can be represented as a $d \times 1$ column vector.

$$|\psi\rangle = \sum_{i=1}^d \langle e_i | \psi \rangle |e_i\rangle \sim \begin{pmatrix} \langle e_1 | \psi \rangle \\ \langle e_2 | \psi \rangle \\ \vdots \\ \langle e_d | \psi \rangle \end{pmatrix}.$$

Similarly, a bra $\langle\psi| \in \mathcal{H}^*$ can be represented as a $1 \times d$ row vector.

$$\langle\psi| = \sum_{i=1}^d \langle\psi|e_i\rangle\langle e_i| \sim \left(\langle\psi|e_1\rangle \quad \langle\psi|e_2\rangle \quad \dots \quad \langle\psi|e_d\rangle \right).$$

Direct Sum Spaces

Given two Hilbert spaces \mathcal{H} and \mathcal{H}' , we introduce the *direct sum* Hilbert space of \mathcal{H} and \mathcal{H}' , denoted $\mathcal{H} \oplus \mathcal{H}'$. This space consists of linear combinations of tuples in $\mathcal{H} \times \mathcal{H}'$, which we denote by $|\psi\rangle \oplus |\psi'\rangle$, where $|\psi\rangle \in \mathcal{H}$ and $|\psi'\rangle \in \mathcal{H}'$. The direct sum space is motivated via its inner product, which we want to be a sesquilinear extension of the relation

$$\langle |\psi\rangle \oplus |\psi'\rangle, |\varphi\rangle \oplus |\varphi'\rangle \rangle = \langle\psi|\varphi\rangle + \langle\psi'|\varphi'\rangle$$

for any two tuples $|\psi\rangle \oplus |\psi'\rangle, |\varphi\rangle \oplus |\varphi'\rangle \in \mathcal{H} \oplus \mathcal{H}'$. This inner product is positive-definite if and only if $\alpha(|\psi\rangle \oplus |\psi'\rangle) + |\varphi\rangle \oplus |\varphi'\rangle = (\alpha|\psi\rangle + |\varphi\rangle) \oplus (\alpha|\psi'\rangle + |\varphi'\rangle)$ for any $\alpha \in \mathbb{C}$. This constitutes the rule for multiplication by a scalar and addition of elements in $\mathcal{H} \oplus \mathcal{H}'$ and we may thus write

$$\mathcal{H} \oplus \mathcal{H}' := \text{span} \{ |\psi\rangle \oplus |\psi'\rangle : |\psi\rangle \in \mathcal{H}, |\psi'\rangle \in \mathcal{H}' \}.$$

An orthonormal basis of this space is given by $\{|e_i\rangle \oplus \mathbf{0}\} \cup \{\mathbf{0} \oplus |e'_j\rangle\}$, where $\{|e_i\rangle\}$ is an orthonormal basis of \mathcal{H} and $\{|e'_j\rangle\}$ is an orthonormal basis of \mathcal{H}' . Hence, $\dim\{\mathcal{H} \oplus \mathcal{H}'\} = d + d'$, where d and d' are the Hilbert dimensions of \mathcal{H} and \mathcal{H}' , respectively.

Tensor Product Spaces

Given two Hilbert spaces \mathcal{H} and \mathcal{H}' , we define the *tensor product* Hilbert space of \mathcal{H} and \mathcal{H}' , denoted $\mathcal{H} \otimes \mathcal{H}'$. The space consists of linear combinations of tuples (pure tensors) in $\mathcal{H} \times \mathcal{H}'$, denoted $|\psi\rangle \otimes |\psi'\rangle$, where $|\psi\rangle \in \mathcal{H}$

2. PRELIMINARIES

and $|\psi'\rangle \in \mathcal{H}'$. Again, we motivate the tensor product space via its inner product, which we want to be a sesquilinear extension of the relation

$$\langle \langle \psi | \otimes \langle \psi' |, |\varphi\rangle \otimes |\varphi'\rangle \rangle = \langle \psi | \varphi \rangle \langle \psi' | \varphi' \rangle.$$

for any two tuples $|\psi\rangle \otimes |\psi'\rangle, |\varphi\rangle \otimes |\varphi'\rangle \in \mathcal{H} \otimes \mathcal{H}'$. This inner product is positive-definite if and only if the following relations are satisfied. For any $\alpha \in \mathbb{C}$, we need

$$\begin{aligned} \alpha(|\psi\rangle \otimes |\psi'\rangle) &= (\alpha|\psi\rangle) \otimes |\psi'\rangle = |\psi\rangle \otimes (\alpha|\psi'\rangle), \\ |\psi\rangle \otimes |\psi'\rangle + |\psi\rangle \otimes |\varphi'\rangle &= |\psi\rangle \otimes (|\psi'\rangle + |\varphi'\rangle) \quad \text{and} \\ |\psi\rangle \otimes |\psi'\rangle + |\varphi\rangle \otimes |\psi'\rangle &= (|\psi\rangle + |\varphi\rangle) \otimes |\psi'\rangle. \end{aligned}$$

The tensor product Hilbert space is then defined as the vector space built from linear combinations of all pure tensors modulus the above equivalence relations. Moreover, if \mathcal{H} has an orthonormal basis $\{|e_i\rangle\}_i$ and \mathcal{H}' has an orthonormal basis $\{|e'_j\rangle\}_j$, then the pure tensors $|e_i\rangle \otimes |e'_j\rangle \in \mathcal{H} \otimes \mathcal{H}'$ form an orthonormal basis of $\mathcal{H} \otimes \mathcal{H}'$ and $\dim \{\mathcal{H} \otimes \mathcal{H}'\} = d \cdot d'$.

2.1.2 Operators on Hilbert Spaces

Linear Operators

We denote the set of *linear operators* from \mathcal{H} to \mathcal{H}' (vector space *homomorphisms*) by $\mathcal{L}(\mathcal{H}, \mathcal{H}')$. Every operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ has a unique decomposition into any pair of orthonormal bases, $\{|e_i\rangle\}$ of \mathcal{H} and $\{|e'_j\rangle\}$ of \mathcal{H}' . The operator can be represented as a $d' \times d$ matrix in these bases.

$$L = \sum_{i,j} \langle e'_j | L | e_i \rangle | e'_j \rangle \langle e_i | \sim \begin{pmatrix} [L]_{11} & [L]_{12} & [L]_{13} & \cdots & [L]_{1d} \\ [L]_{21} & [L]_{22} & [L]_{23} & \cdot & \vdots \\ \vdots & \cdot & \cdot & \cdot & \vdots \\ [L]_{d'1} & \cdots & \cdots & \cdots & [L]_{d'd} \end{pmatrix}, \quad (2.2)$$

where $[L]_{ji} := \langle e'_j | L | e_i \rangle$. Note that representations of kets, bras and operators in a particular basis are only used as illustrations throughout this thesis.

For every $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, we define its *adjoint* operator $L^\dagger \in \mathcal{L}(\mathcal{H}', \mathcal{H})$ as the unique operator that satisfies

$$\langle \psi | L | \varphi \rangle = \overline{\langle \varphi | L^\dagger | \psi \rangle} \quad \text{for all } |\varphi\rangle \in \mathcal{H} \text{ and } |\psi\rangle \in \mathcal{H}'.$$

(This is equivalent to the condition $\langle |\psi\rangle, L|\varphi\rangle \rangle = \langle L^\dagger|\psi\rangle, |\varphi\rangle \rangle$ expressed in terms of the inner product.) In particular, this definition, together with (2.1), implies that a ket $L|\varphi\rangle$ has the dual bra $\langle \varphi | L^\dagger$.

2.1 Linear Algebra on Hilbert Spaces

Moreover, we will also need the *transpose* L^T of L with regards to a pair of bases, $\{|e_i\rangle\}$ of \mathcal{H} and $\{|e'_j\rangle\}$ of \mathcal{H}' . This is defined as

$$L^T := \sum_{i,j} \langle e'_j | L | e_i \rangle | e_i \rangle \langle e'_j |.$$

The *support* of an operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ is the subspace of \mathcal{H} spanned by all kets that are not mapped to zero by L , that is

$$\text{supp } \{L\} := \text{span } \{|\varphi\rangle : |\varphi\rangle \in \mathcal{H} \text{ and } L|\varphi\rangle \neq 0\}.$$

The *rank* of L is the dimension of its support, $\text{rank } \{L\} := \dim \{\text{supp } \{L\}\}$. In contrast, the *kernel* of L is the subspace orthogonal to the support of L , namely

$$\text{kern } \{L\} := \text{span } \{|\varphi\rangle : |\varphi\rangle \in \mathcal{H} \text{ and } L|\varphi\rangle = 0\} = \mathcal{H} / \text{supp } \{L\}.$$

Finally, the *image* of L is the subspace of \mathcal{H}' spanned by L , that is

$$\text{image } \{L\} := \text{span } \{L|\varphi\rangle : |\varphi\rangle \in \mathcal{H}\}.$$

Projectors, Identity and Inverse

The linear operators from \mathcal{H} onto itself (vector space *endomorphisms*) are denoted $\mathcal{L}(\mathcal{H}) := \mathcal{L}(\mathcal{H}, \mathcal{H})$. A *projector* into a subspace \mathcal{H}' of \mathcal{H} is an operator $P \in \mathcal{L}(\mathcal{H})$ with $\text{supp } \{P\} = \text{image } \{P\} = \mathcal{H}'$ that acts as an *identity* on all $|\varphi'\rangle \in \mathcal{H}'$, i.e. $P|\varphi'\rangle = |\varphi'\rangle$. Thus, $P^2 = P$ and $P = P^\dagger$. We use the symbol $\mathbb{1}$ to denote the *identity operator* on \mathcal{H} , which is the projector from \mathcal{H} onto \mathcal{H} . The identity operator can be decomposed in any orthonormal basis $\{|e_i\rangle\}$ of \mathcal{H} as $\mathbb{1} = \sum_i |e_i\rangle \langle e_i|$. For any linear operator $L \in \mathcal{L}(\mathcal{H})$, we denote the projector onto its support by Π^L .

For operators $L \in \mathcal{L}(\mathcal{H})$, we define their *inverse* (if it exists), $L^{-1} \in \mathcal{L}(\mathcal{H})$, as the unique operator satisfying $L^{-1}L = LL^{-1} = \mathbb{1}$. We will often use a *generalized inverse*, which is defined for every operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ and is the inverse of L on its support. This means, we define the generalized inverse L^{-1} as the unique operator satisfying $L^{-1}L = \Pi^L$ and $\text{supp } \{L^{-1}\} = \text{image } \{L\}$.

Isomorphisms, Unitaries and Isometries

An *isomorphism* is a linear bijective map associating elements of two structured sets in a way that preserves that structure. In particular, an isomorphism $u : \mathcal{H} \rightarrow \mathcal{H}'$ between Hilbert spaces \mathcal{H} and \mathcal{H}' preserves the inner product and, thus, satisfies

$$\langle |\psi\rangle, |\varphi\rangle \rangle = \langle u(|\psi\rangle), u(|\varphi\rangle) \rangle \quad (2.3)$$

2. PRELIMINARIES

for all $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$. We call two Hilbert spaces \mathcal{H} and \mathcal{H}' isomorphic, denoted $\mathcal{H} \cong \mathcal{H}'$, if there exists an isomorphism between \mathcal{H} and \mathcal{H}' . Two Hilbert spaces are isomorphic if and only if they have the same Hilbert dimension. (To see this, note that an isomorphism can always be seen as a bijective map between elements of two orthonormal bases.)

A *unitary operator* is an operator $U \in \mathcal{L}(\mathcal{H})$ that is an isomorphism between \mathcal{H} and itself. Since an isomorphism has to satisfy (2.3), we have $\langle\psi|\varphi\rangle = \langle\psi|U^\dagger U|\varphi\rangle$ or, equivalently, $U^\dagger U = \mathbb{1}$. Hence, the inverse of a unitary operator U is well defined and $U^{-1} = U^\dagger$.

A *partial isometry* is an operator $V \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ that is an isomorphism between $\text{supp}\{V\} \subseteq \mathcal{H}$ and $\text{image}\{V\} \subseteq \mathcal{H}'$. A partial isometry satisfies $V^\dagger V = \Pi^V$ and, thus, $V^{-1} = V^\dagger$ is its generalized inverse (and a partial isometry too). An *isometry* is a partial isometry with full support on \mathcal{H} .

In the following, we denote the set of unitary operators on a Hilbert space \mathcal{H} as $\mathcal{U}(\mathcal{H})$ and the set of partial isometries from \mathcal{H} to \mathcal{H}' as $\mathcal{U}(\mathcal{H}, \mathcal{H}')$. Isometries can always be understood as *embeddings*. Let \mathcal{H} and \mathcal{H}' be two Hilbert spaces such that $\dim\{\mathcal{H}\} < \dim\{\mathcal{H}'\}$ and let $V \in \mathcal{U}(\mathcal{H}, \mathcal{H}')$ be an isometry that embeds \mathcal{H} into \mathcal{H}' , i.e. it satisfies $V^\dagger V = \mathbb{1}$ on \mathcal{H} . Then, for every operator $L \in \mathcal{L}(\mathcal{H})$, we implicitly define its embedding $L' := V L V^\dagger \in \mathcal{L}(\mathcal{H}')$.

Trace

For any Hilbert space \mathcal{H} , the *trace* of an operator in $\mathcal{L}(\mathcal{H})$ is the (linear) functional $\text{tr} : \mathcal{L}(\mathcal{H}) \rightarrow \mathbb{C}$ with the defining properties

$$\text{tr}(AB) = \text{tr}(BA) \quad \text{and} \quad \text{tr}(\mathbb{1}) = d \quad (2.4)$$

for all linear operators $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ and $B \in \mathcal{L}(\mathcal{H}', \mathcal{H})$. In particular the trace is invariant under unitary conjugation, $\text{tr}(U^\dagger L U) = \text{tr}(L)$ for any $U \in \mathcal{U}(\mathcal{H})$ and $L \in \mathcal{L}(\mathcal{H})$. This implies that there exists a representation of the trace as a functional on matrix representations that is independent of the choice of basis used for the representation.

If we choose an orthonormal basis $\{|e_i\rangle\}$ for \mathcal{H} and $\{|e'_j\rangle\}$ for \mathcal{H}' , the operators AB and BA can be represented as matrices with entries

$$[AB]_{ji} = \sum_k \langle e'_j | A | e_k \rangle \langle e_k | B | e'_i \rangle \quad \text{and} \quad [BA]_{ji} = \sum_k \langle e'_k | A | e_i \rangle \langle e_j | B | e'_k \rangle$$

and it is easy to verify that the only functional that satisfies (2.4) is the sum of the diagonal elements, $\sum_i [AB]_{ii} = \sum_i [BA]_{ii}$. Hence, for any $L \in \mathcal{L}(\mathcal{H})$ and any basis $\{|e_i\rangle\}$ of \mathcal{H} , we have $\text{tr}(L) = \sum_i [L]_{ii} = \sum_i \langle e_i | L | e_i \rangle$, which is how the trace of a matrix is commonly defined.

Singular Values

For any operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, there always exists a pair of bases $\{|e_i\rangle\}$ of \mathcal{H} and $\{|e'_i\rangle\}$ of \mathcal{H}' such that L can be decomposed as

$$L = \sum_i s_i |e'_i\rangle\langle e_i| \quad \text{and} \quad s_1 \geq s_2 \geq \dots \geq s_m > 0, \quad (2.5)$$

where m is the rank of L . This is called the *singular value decomposition* and the unique positive $s_i = s_i(L)$ are called *singular values*. The singular values are invariant under unitary rotations, as these operations can be absorbed into the basis. Thus, for any $U \in \mathcal{U}(\mathcal{H})$ and $V' \in \mathcal{U}(\mathcal{H}')$,

$$s_i(V'LU) = s_i(L). \quad (2.6)$$

Support and image of L can be expressed in terms of the two bases, that is $\text{supp } \{L\} = \text{span } \{|e_i\rangle\}$ and $\text{image } \{L\} = \text{span } \{|e'_i\rangle\}$.

Note that the operator $L^\dagger L = \sum_i s_i^2 |e_i\rangle\langle e_i| \in \mathcal{L}(\mathcal{H})$ has a unique positive square root, namely the *modulus* of L ,

$$|L| := \sqrt{L^\dagger L} = \sum_i s_i |e_i\rangle\langle e_i|.$$

Comparing this with the singular value decomposition (2.5), we find the *polar decomposition*, $L = W|L|$, where $W : |e_i\rangle \mapsto |e'_i\rangle$ is a partial isometry from \mathcal{H} to \mathcal{H}' defined through the bases of the singular value decomposition of L .

Hilbert-Schmidt Inner Product and Schmidt Decomposition

Given two Hilbert spaces \mathcal{H} with orthonormal basis $\mathcal{B} = \{|e_i\rangle\}$ and \mathcal{H}' , we associate operators in $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ with tensors in $\mathcal{H} \otimes \mathcal{H}'$ through the *isomorphism* $\text{vec}_{\mathcal{B}} : L \mapsto \sum_i |e_i\rangle \otimes L|e_i\rangle$. Using the decomposition of L in (2.2), we find that $\text{vec}(L)$ can be written as

$$\text{vec}_{\mathcal{B}}(L) = \sum_{i,j} [L]_{j,i} |e_i\rangle \otimes |e'_j\rangle,$$

which simply corresponds to the rearrangement of the matrix entries of L (in the given bases) as a column vector.

This isomorphism is useful because it induces an inner product on linear operators. We define the *Hilbert-Schmidt inner product* on the complex Hilbert space $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ as

$$\langle A, B \rangle := \langle \text{vec}_{\mathcal{B}}(A), \text{vec}_{\mathcal{B}}(B) \rangle = \sum_{i,j} \langle e_j | e_i \rangle \langle e_j | A^\dagger B | e_i \rangle = \sum_i \langle e_i | A^\dagger B | e_i \rangle,$$

2. PRELIMINARIES

for any $A, B \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$. This expression is equal to the trace and, thus, independent of the basis \mathcal{B} chosen for the isomorphism:

$$\langle A, B \rangle = \text{tr}(A^\dagger B). \quad (2.7)$$

The tensor $\text{vec}_{\mathcal{B}}(\mathbb{1}) = \sum_i |e_i\rangle \otimes |e_i\rangle$ takes a special role in the analysis of quantum systems. In particular, we will need the following property.

Lemma 2.1 (Mirror Lemma). *Let \mathcal{H} be a Hilbert space with orthonormal basis \mathcal{B} and $L \in \mathcal{L}(\mathcal{H})$, then $(\mathbb{1} \otimes L) \text{vec}_{\mathcal{B}}(\mathbb{1}) = (L^\top \otimes \mathbb{1}) \text{vec}_{\mathcal{B}}(\mathbb{1})$, where the transpose is taken with regards to \mathcal{B} .*

Proof. By inspection. \square

This, together with the above isomorphism, can be used to prove the existence of the *Schmidt decomposition*.

Lemma 2.2 (Schmidt Decomposition). *Let $\mathcal{H}, \mathcal{H}'$ be Hilbert spaces and let $|\theta\rangle \in \mathcal{H} \otimes \mathcal{H}'$. Then, there exist orthonormal bases $\{|\vartheta_i\rangle\}$ of \mathcal{H} and $\{|\vartheta'_i\rangle\}$ of \mathcal{H}' as well as non-negative numbers s_i such that*

$$|\theta\rangle = \sum_i s_i |\vartheta_i\rangle \otimes |\vartheta'_i\rangle.$$

Proof. Given a linear operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ and an orthonormal basis $\mathcal{B} = \{|e_i\rangle\}$, we use the singular value decomposition of L to get

$$\text{vec}_{\mathcal{B}}(L) = \text{vec}_{\mathcal{B}}(V^\dagger S U) = \sum_i s_i(L) U^\top |e_i\rangle \otimes V^\dagger |e_i\rangle, \quad (2.8)$$

where $S = \sum_i s_i(L) |e_i\rangle\langle e_i|$ is diagonal in \mathcal{B} . The isometries $U \in \mathcal{U}(\mathcal{H})$ and $V \in \mathcal{U}(\mathcal{H}', \mathcal{H})$ map the bases of the singular value decomposition to \mathcal{B} . Hence, every tensor can be written in the form (2.8), which concludes the proof. \square

Operator Norms

The singular value decomposition allows the definition of various norms on $\mathcal{L}(\mathcal{H}, \mathcal{H}')$. In particular, we will often use the *Schatten norms*. For $p \geq 1$, they are defined as (cf. e.g. [Bha97, Wat08])

$$\|L\|_p := \left(\sum_{i=1}^m s_i(L)^p \right)^{\frac{1}{p}}.$$

In particular, we will often use the infinity norm, $\|\cdot\|_\infty$, which is equal to the induced norm of the underlying Hilbert space norm, that is

$$\|L\|_\infty = s_1(L) = \sup_{|\psi\rangle \in \mathcal{H}} \frac{\|L|\psi\rangle\|}{\| |\psi\rangle \|}.$$

The 1-norm is equal to the trace norm,

$$\|L\|_1 = \sum_{i=1}^m s_i(L) = \text{tr } |L|.$$

And, finally, the 2-norm is equal to the induced norm of the Hilbert-Schmidt inner product,

$$\|L\|_2 = \left(\sum_{i=1}^m s_i(L)^2 \right)^{1/2} = \sqrt{\langle L, L \rangle}.$$

These three norms satisfy $\|L\|_\infty \leq \|L\|_2 \leq \|L\|_1$. And, thanks to (2.6), they are *invariant under unitary rotations*. Hence, $\|L\| = \|V'LU\|$, where $\|\cdot\|$ denotes any of the Schatten norms introduced above. For any three operators $L, M, N \in \mathcal{L}(\mathcal{H})$, these norms satisfy

$$\|MLN\| \leq \|M\|_\infty \|L\| \|N\|_\infty. \quad (2.9)$$

Moreover, these norms are *sub-multiplicative*, i.e. $\|MN\| \leq \|M\| \|N\|$.

2.1.3 Positive Semi-Definite Operators

Hermitian Operators

An operator $M \in \mathcal{L}(\mathcal{H})$ is called *self-adjoint* or *Hermitian* if $M^\dagger = M$. The set of Hermitian operators on \mathcal{H} is denoted $\mathcal{H}(\mathcal{H}) := \{M \in \mathcal{L}(\mathcal{H}) : M^\dagger = M\}$. Note that the real vector space $\mathcal{H}(\mathcal{H})$ together with the Hilbert-Schmidt inner product in (2.7) form a *real Hilbert space*.

Every $M \in \mathcal{H}(\mathcal{H})$ has an *eigenvalue decomposition*, namely

$$M = \sum_i \lambda_i |e_i\rangle\langle e_i|, \quad \text{where } \lambda_i \in \mathbb{R} \text{ and } |\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_m| > 0.$$

The coefficients $\lambda_i = \lambda_i(M)$ are called *eigenvalues* of M and $|e_i\rangle$ are *eigenvectors*. Together, the eigenvectors form an *eigenbasis* $\{|e_i\rangle\}$, which is an orthonormal basis of $\text{supp } \{M\} = \text{image } \{M\}$. This eigenbasis is unique if and only if all the eigenvalues are mutually different. The singular values can be expressed in terms of the eigenvalues, that is $s_i(M) = |\lambda_i(M)|$.¹

The eigenvalue decomposition is particularly useful to define the action of functions on operators. Note, for example, that $M^2 = MM =$

¹The existence of the eigenvalue decomposition follows, for example, from the fact that the singular value decompositions of $M^\dagger M = \sum_i s_i^2 |e_i\rangle\langle e_i|$ and $MM^\dagger = \sum_i s_i^2 |e'_i\rangle\langle e'_i|$ have to be the same. Thus, if all singular values are mutually different, we have $|e'_i\rangle = e^{i\varphi} |e_i\rangle$ and the eigenvalue decomposition is of the form $M = \sum_i \pm s_i |e_i\rangle\langle e_i|$. More generally, an eigenvalue decomposition exists for all *normal* $L \in \mathcal{L}(\mathcal{H})$, where normal means that $L^\dagger L = LL^\dagger$.

2. PRELIMINARIES

$\sum_i \lambda_i^2 |e_i\rangle\langle e_i|$, $|M| = \sum_i |\lambda_i| |e_i\rangle\langle e_i|$ and $M^{-1} = \sum_i \frac{1}{\lambda_i} |e_i\rangle\langle e_i|$. We generalize this and define the action of arbitrary functions $f : \mathbb{R}/\{0\} \rightarrow \mathbb{R}$ on Hermitian operators as

$$f : M = \sum_i \lambda_i |e_i\rangle\langle e_i| \mapsto \sum_i f(\lambda_i) |e_i\rangle\langle e_i|.$$

Note that the function f only acts on the support of the Hermitian operator, leaving its kernel intact.

For any $M \in \mathcal{H}(\mathcal{H})$, we denote by $\{M\}_+$ and $\{M\}_-$ the projections of M onto its positive and negative eigenspace, respectively, that is

$$\{M\}_+ = \sum_{i:\lambda_i>0} \lambda_i |e_i\rangle\langle e_i| \quad \text{and} \quad \{M\}_- = \sum_{i:\lambda_i<0} \lambda_i |e_i\rangle\langle e_i|.$$

Positive Semi-Definite Operators

The set of *positive semi-definite* operators, $\mathcal{P}(\mathcal{H}) \subset \mathcal{H}(\mathcal{H})$ is the set of operators that have positive inner products with all vectors in \mathcal{H} , namely

$$\begin{aligned} \mathcal{P}(\mathcal{H}) &:= \{M \in \mathcal{H}(\mathcal{H}) : \langle \psi | M | \psi \rangle \geq 0 \text{ for all } |\psi\rangle \in \mathcal{H}\} \\ &= \{M \in \mathcal{H}(\mathcal{H}) : \lambda_i(M) > 0 \ \forall i\}. \end{aligned}$$

We write $M \geq 0$ if and only if $M \in \mathcal{P}(\mathcal{H})$. In the following, we simply call these operators *positive*. Moreover, given two Hermitian operators $M, N \in \mathcal{H}(\mathcal{H})$, we write $M \geq N$ if and only if $M - N \geq 0$ and we write $M \leq N$ if and only if $N - M \geq 0$. Note that this relation constitutes a partial order on $\mathcal{H}(\mathcal{H})$.

2.1.4 Operators on Tensor Spaces

We will often use linear operators on *tensor product* spaces, for example $\mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \bar{\mathcal{H}} \otimes \bar{\mathcal{H}}') \cong \mathcal{L}(\mathcal{H}, \bar{\mathcal{H}}) \otimes \mathcal{L}(\mathcal{H}', \bar{\mathcal{H}}')$, where the second isomorphic space is constructed from the Hilbert spaces $\mathcal{L}(\mathcal{H}, \bar{\mathcal{H}})$ and $\mathcal{L}(\mathcal{H}', \bar{\mathcal{H}}')$ using (??). Hence, we can write every element $L \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \bar{\mathcal{H}} \otimes \bar{\mathcal{H}}')$ as a linear combination of the form

$$L = \sum_{\alpha, \beta} \chi_{\alpha\beta} \sigma_\alpha \otimes \sigma'_\beta, \quad \text{where } \chi_{\alpha\beta} \in \mathbb{C}, \quad (2.10)$$

and $\{\sigma_\alpha\}$ and $\{\sigma'_\beta\}$ are bases of $\mathcal{L}(\mathcal{H}, \bar{\mathcal{H}})$ and $\mathcal{L}(\mathcal{H}', \bar{\mathcal{H}}')$, respectively. If L is Hermitian, then the $\chi_{\alpha\beta}$ can be chosen real and the bases can be chosen Hermitian. (This follows directly from the fact that $\mathcal{H}(\mathcal{H})$ is a real Hilbert space.) Furthermore, this allows us to introduce a trivial extension of operators $L \in \mathcal{L}(\mathcal{H}, \bar{\mathcal{H}})$ to operators $L \otimes \mathbb{1}$ in $\mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \bar{\mathcal{H}} \otimes \bar{\mathcal{H}}')$. In the following, this extension is assumed implicitly whenever an operator defined on a subspace is applied to a tensor space.

2.1.5 Completely Positive Maps

Super-Operators

A *super-operator* is a linear map from linear operators on one Hilbert space to linear operators on another Hilbert space. For example, if the two Hilbert spaces are \mathcal{H} and \mathcal{H}' , we denote the set of super-operators from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H}')$ by $\mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$. The super-operators form a vector space and, since every operator in $\mathcal{L}(\mathcal{H})$ can be decomposed according to (2.10), the action of a super-operator on an operator in a (larger) tensor space is well-defined via linearity. We use the symbol \circ to denote concatenations of super-operators, for example $(\mathcal{E} \circ \mathcal{F})[\cdot] = \mathcal{E}[\mathcal{F}[\cdot]]$.

For any super-operator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$, its adjoint super-operator, $\mathcal{E}^\dagger \in \mathcal{L}(\mathcal{L}(\mathcal{H}'), \mathcal{L}(\mathcal{H}))$, is defined via the Hilbert-Schmidt inner product as the unique operator satisfying

$$\langle \mathcal{E}[A], B \rangle = \langle A, \mathcal{E}^\dagger[B] \rangle \quad \text{for all } A \in \mathcal{L}(\mathcal{H}), B \in \mathcal{L}(\mathcal{H}').$$

Completely Positive Maps

super-operators that (consistently) map positive operators onto positive operators are called *completely positive maps* (CPMs).

Definition 2.2 (Completely Positive Map). *Let \mathcal{H} and \mathcal{H}' be Hilbert spaces. A super-operator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ is called completely positive, if, for any auxiliary Hilbert space \mathcal{H}'' , it holds that*

$$\mathcal{E}[M] \geq 0 \quad \text{for all } M \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H}').$$

An example of such a completely positive super-operator is the *conjugation* with an operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, that is the map $\mathcal{L} : M \mapsto LML^\dagger$. We will often use the following basic property of completely positive maps. Let $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ be completely positive, then

$$A \geq B \implies \mathcal{E}[A] \geq \mathcal{E}[B] \quad \text{for all } A, B \in \mathcal{H}(\mathcal{H}). \quad (2.11)$$

Furthermore, we take note of the following property of positive semi-definite operators. For any $M, N \in \mathcal{P}(\mathcal{H})$, we have

$$\text{tr}(MN) = \text{tr}(\sqrt{M}N\sqrt{M}) \geq 0,$$

where the last inequality follows from the fact that the conjugation with \sqrt{M} is a completely positive map. In particular, if $X, Y \in \mathcal{H}(\mathcal{H})$ satisfy $X \geq Y$, we find $\text{tr}(MX) \geq \text{tr}(MY)$.

A completely positive map \mathcal{E} is called *trace preserving* if $\text{tr}(\mathcal{E}[M]) = \text{tr}(M)$ for all $M \in \mathcal{L}(\mathcal{H})$. We denote the set of all trace preserving completely positive maps (*TP-CPMs*) from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H}')$ by $\mathcal{T}(\mathcal{H}, \mathcal{H}')$. If \mathcal{E}

2. PRELIMINARIES

is a TP-CPM, then its adjoint \mathcal{E}^\dagger is completely positive and *unital*. Since

$$\text{tr}(M) = \text{tr}(\mathcal{E}[M]) = \text{tr}(\mathcal{E}^\dagger[\mathbb{1}]M) \quad \text{for all } M \in \mathcal{L}(\mathcal{H}),$$

unital maps are defined by the property that they map the identity onto an identity, i.e. $\mathcal{E}^\dagger[\mathbb{1}] = \mathbb{1}$.

Finally, a completely positive map \mathcal{E} is called *trace non-increasing* if $\text{tr}(\mathcal{E}[M]) \leq \text{tr}(M)$ for all $M \in \mathcal{P}(\mathcal{H})$. It is easy to verify (using the same argument as above) that its adjoint is completely positive and *sub-unital*, i.e. it satisfies $\mathcal{E}^\dagger[\mathbb{1}] \leq \mathbb{1}$.

Partial Trace

Given a bipartite Hilbert space $\mathcal{H} \otimes \mathcal{H}'$, we are interested in the *partial trace* super-operator, denoted $\text{tr}_{\mathcal{H}'} \in \mathcal{T}(\mathcal{H} \otimes \mathcal{H}', \mathcal{H})$. The partial trace is defined as the adjoint super-operator to $L \mapsto L \otimes \mathbb{1}$, which maps operators $L \in \mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H} \otimes \mathcal{H}')$. It is easy to verify, using (2.7), that the *trace* is the special case of the partial trace where $\mathcal{H} \cong \mathbb{C}$.

To justify this definition, let us investigate the action of the partial trace $\text{tr}_{\mathcal{H}'}$ of a product state $K \otimes L$, where $K \in \mathcal{L}(\mathcal{H})$ and $L \in \mathcal{L}(\mathcal{H}')$, on an arbitrary state $Q \in \mathcal{L}(\mathcal{H})$. We have

$$\langle Q, \text{tr}_{\mathcal{H}'}(K \otimes L) \rangle = \langle K \otimes L, Q \otimes \mathbb{1}_{\mathcal{H}'} \rangle = \langle K, Q \rangle \text{tr}(L).$$

Since this holds for all Q , we have $\text{tr}_{\mathcal{H}'}(K \otimes L) = \text{tr}(L)K$.

The partial trace is cyclic in operators on the same Hilbert space and commutes with operators on other Hilbert spaces. Specifically, let $L \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \mathcal{H} \otimes \mathcal{H}'')$ and $K \in \mathcal{L}(\mathcal{H}'', \mathcal{H}')$. We use the decomposition $L = \sum_{\alpha, \beta} \chi_{\alpha\beta} \sigma_\alpha \otimes \sigma'_\beta$ of (2.10). Since the partial trace is linear and K only acts on \mathcal{H}' , we have, using the cyclicity of the trace,

$$\text{tr}_{\mathcal{H}'}(KL) = \sum_{\alpha, \beta} \text{tr}(K \sigma'_\beta) \sigma_\alpha = \sum_{\alpha, \beta} \text{tr}(\sigma'_\beta K) \sigma_\alpha = \text{tr}_{\mathcal{H}'}(LK).$$

Moreover, using this decomposition, it is easy to verify that $\text{tr}_{\mathcal{H}}(KL) = K \text{tr}_{\mathcal{H}}(L)$ and $\text{tr}_{\mathcal{H}}(LK) = \text{tr}_{\mathcal{H}}(L)K$.

Choi-Jamiołkowski Isomorphism

In the same spirit as the vec-isomorphism between operators $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ and tensors $\mathcal{H} \otimes \mathcal{H}'$, we now define the *Choi-Jamiołkowski isomorphism* [Jam72], Ω , between super-operators in $\mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ and operators in $\mathcal{L}(\mathcal{H}' \otimes \mathcal{H})$.

$$\Omega : \mathcal{E} \mapsto \omega^\mathcal{E} = \mathcal{E}[|\Gamma\rangle\langle\Gamma|], \quad \text{where } |\Gamma\rangle = \sum_i |e_i\rangle \otimes |e_i\rangle \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H})$$

and $\{|e_i\rangle\}$ is an orthonormal basis of \mathcal{H} . The state $\omega^\mathcal{E}$ is called the Choi-Jamiolkowski state of \mathcal{E} . The inverse operation Ω^{-1} maps Choi-Jamiolkowski states $\omega^\mathcal{E} \in \mathcal{L}(\mathcal{H}' \otimes \mathcal{H})$ to super-operators

$$\mathcal{E} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}'), \quad X \mapsto \text{tr}_{\mathcal{H}}(\omega^\mathcal{E}(\mathbb{1}_{\mathcal{H}'} \otimes X^T))$$

where the transpose is taken with regards to the basis $\{|e_i\rangle\}$.

There are various relations between properties of super-operators and properties of the corresponding Choi-Jamiolkowski states. The following can be verified by inspection.

- \mathcal{E} is completely positive $\iff \omega^\mathcal{E}$ is positive semi-definite.
- \mathcal{E} is a trace-preserving $\iff \text{tr}_{\mathcal{H}'}(\omega^\mathcal{E}) = \mathbb{1}_{\mathcal{H}}$.
- \mathcal{E} is unital $\iff \text{tr}_{\mathcal{H}}(\omega^\mathcal{E}) = \mathbb{1}_{\mathcal{H}'}$.

Kraus Operators and Stinespring Dilation

The following two lemmas are of crucial importance in quantum information theory. They describe two alternative representations of completely positive¹ maps, especially trace non-increasing and trace preserving maps.

Every completely positive super-operator can be represented as a sum of conjugations of the input with *Kraus operators*. [HK69, HK70].

Lemma 2.3 (Kraus Representation). *A super-operator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ is completely positive if and only if there exists a finite set of linear operators $\{E_k\}$, $E_k \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ such that*

$$\mathcal{E}[A] = \sum_k E_k A E_k^\dagger \quad \text{for all } A \in \mathcal{L}(\mathcal{H}).$$

Furthermore, a completely positive \mathcal{E} is trace non-increasing if and only if $\sum_k E_k^\dagger E_k \leq \mathbb{1}$ and it is trace preserving if and only if $\sum_k E_k^\dagger E_k = \mathbb{1}$.

The operators $\{E_k\}$ are called Kraus operators. Note that the adjoint \mathcal{E}^\dagger of \mathcal{E} is completely positive and has Kraus operators $\{E_k^\dagger\}$ since

$$\text{tr}(\mathcal{E}^\dagger[B]A) = \text{tr}(B \mathcal{E}[A]) = \text{tr}\left(\sum_k E_k^\dagger B E_k A\right) \quad \text{for all } A \in \mathcal{L}(\mathcal{H}).$$

Moreover, every CPM can be decomposed into its *Stinespring dilation* form [Sti55] as follows.

Lemma 2.4 (Stinespring Dilation). *A super-operator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ is completely positive if and only if there exists a Hilbert space \mathcal{H}'' and an operator $L \in \mathcal{L}(\mathcal{H}, \mathcal{H}' \otimes \mathcal{H}'')$ such that*

$$\mathcal{E}(A) = \text{tr}_{\mathcal{H}''}(L A L^\dagger) \quad \text{for all } A \in \mathcal{L}(\mathcal{H}).$$

Moreover, if \mathcal{E} is trace preserving then L is an isometry. If \mathcal{E} is trace non-increasing, then L is an isometry followed by a projection in $\mathcal{P}(\mathcal{H}'')$.

¹An extension of this to arbitrary super-operators is possible (see, e.g. [Wat08]).

2. PRELIMINARIES

Proofs of these lemmas can be found in most quantum information textbooks (see, e.g. [NC00]). We will only prove the result for trace non-increasing CPMs, which is a bit less standard.

Proof. Let \mathcal{E} be a trace non-increasing CPM. First, note that its Kraus representation has to satisfy, for all $A \in \mathcal{P}(\mathcal{H})$,

$$\mathrm{tr}(\mathcal{E}(A)) = \mathrm{tr}\left(\sum_k E_k A E_k^\dagger\right) = \mathrm{tr}\left(\sum_k E_k^\dagger E_k A\right) \leq \mathrm{tr}(A).$$

This is equivalent to the condition $\sum_{k=1}^n E_k^\dagger E_k \leq \mathbb{1}$. Hence, the map can be extended to a trace preserving CPM by adding another Kraus operator into the sum, e.g.

$$E_{n+1} = \left(\mathbb{1} - \sum_{k=1}^n E_k^\dagger E_k\right)^{\frac{1}{2}}.$$

We now construct a Stinespring dilation for this extended operation. In fact, a possible dilation is $L = \sum_{k=1}^{n+1} E_k \otimes |e_k\rangle$, where $\{|e_k\rangle\}$ is an orthonormal basis of \mathcal{H}'' . This operation is an isometry and a Stinespring dilation of \mathcal{E} can be recovered by projecting onto $\sum_{k=1}^n |e_k\rangle\langle e_k|$ after applying L . \square

2.2 Quantum Mechanics

In this thesis, we will use a mathematical model for *quantum mechanics* — the density operator formalism on finite-dimensional Hilbert spaces — that is restricted to physical systems with a finite dimensional configuration space. This means that continuous observables such as position and momentum need to be considered discretized and bounded. While it is unknown whether such a framework is sufficient to describe all possible correlations between observations of physical systems [SW08]¹, it offers the opportunity to focus on the main physical concepts without getting lost in delicate mathematical arguments. Moreover, the main results of this thesis are independent of the actual *dimension* of the physical systems under consideration and some of the results have already been re-derived in a more general model of quantum mechanics that is based on infinite dimensional Hilbert spaces [FABR11] or (von Neumann) operator algebras of observables [BFS11].

This section is partly inspired by Nielsen and Chuang [NC00] as well as Hardy [Har01] and introduces quantum mechanics from a strictly information theoretic perspective.² The most important deviations from the standard treatment of quantum mechanics are pointed out.

¹More precisely, it is unclear whether all correlations can be approximated to arbitrary precision using the density operator formalism.

²Quantum mechanics is often formulated in a way that highlights its relation to

2.2.1 Systems and States

We use a very abstract notion of *physical systems*, describing them as general purpose information carriers without specifying their actual physical realization. In fact, the only system parameter we consider is its dimension, which, as we will see, corresponds to the dimension of the *Hilbert space* used to describe the system.

The *dimension* of an isolated system is given by the number of mutually perfectly distinguishable states that can be prepared on the system. This assumes perfect (idealized) preparation and measurement equipment. Alternatively, the dimension can be interpreted as the information storage capacity of the system. Very abstractly, an encoder is a map \mathcal{E} from a set \mathcal{X} to states of the system and a decoder \mathcal{D} is a map from states of the system to \mathcal{X} . A pair $\{\mathcal{E}, \mathcal{D}\}$ of encoder and decoder is perfect if $x = \mathcal{D}[\mathcal{E}[x]]$ for all $x \in \mathcal{X}$. The dimension of a system is then given by the maximum cardinality of a set \mathcal{X} such that there exist perfect encoders and decoders between \mathcal{X} and states of the system.

The simplest such system, a *qubit* or two-level system, has dimension two and may have different physical realizations. For example, the information could be encoded as the spin degree of freedom of an electron or the polarization of a photon.

Postulate 1 (State Space). *Isolated quantum systems are modeled as Hilbert spaces with the dimension of the system. The system's state is represented by a positive semi-definite operator with unit trace acting on this space.*

The state of an isolated quantum system A is thus fully characterized by all linear functionals on the state space, i.e. the functionals $\text{tr}(\rho_A L_A)$ where $L_A \in \mathcal{L}(\mathcal{H}_A)$ and ρ_A is the state of the system.

We denote isolated physical systems with capital letters, i.e. A, B, C , and their associated Hilbert spaces with $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$. The states themselves are denoted using lowercase greek letters, i.e. ρ_A, σ_B, τ_C , where the subscripts indicate which system is meant when necessary. The dimension of a system A is denoted by $d_A := \dim \{\mathcal{H}_A\}$.

We use $\mathcal{S}_=(\mathcal{H}_A)$ to denote the set of quantum states on A , i.e. the set $\mathcal{S}_=(\mathcal{H}_A) := \{\rho \in \mathcal{P}(\mathcal{H}_A) : \text{tr}(\rho) = 1\}$. Hence, a state of a system A is represented as an element $\rho_A \in \mathcal{S}_=(\mathcal{H}_A)$. In addition, we sometimes employ a larger set of states on \mathcal{H}_A , the sub-normalized states $\mathcal{S}_\leq(\mathcal{H}_A) := \{\rho \in \mathcal{P}(\mathcal{H}_A) : 0 < \text{tr}(\rho) \leq 1\}$. These states by themselves do not have a physical interpretation and usually only appear in technical statements

classical mechanics. In quantum information science, due to its interdisciplinary nature between physics, computer science and information theory, a different approach to quantum mechanics has proven fruitful.

2. PRELIMINARIES

However, they can be seen as normalized quantum states on a Hilbert space $\mathcal{H}_{A'}$ that are projected onto a subspace \mathcal{H}_A of $\mathcal{H}_{A'}$.

We call a positive semi-definite operator $\rho_A \in \mathcal{P}(\mathcal{H}_A)$ *pure* if it has rank one, i.e. if $\text{rank}\{\rho_A\} = 1$. Pure operators can be represented as kets $|\rho\rangle_A \in \mathcal{H}_A$, where $|\rho\rangle_A$ is determined by the relation $\rho_A = |\rho\rangle\langle\rho|_A$ up to a phase factor. An operator that is not pure is called *mixed*. We often employ the completely mixed state on a system A , which is given by $\pi_A := \mathbb{1}_A/d_A$.

In most introductions to quantum mechanics — supposedly for historical reasons — the theory is first formulated in terms of pure quantum states and then later extended to arbitrary mixed states. From an information theoretic perspective, however, mixed states are more fundamental as they are generalizations of classical *random variables*.

Classical Registers

Discrete *probability distributions* can be conveniently represented as states of a quantum system. We call these systems *registers* and typically denote them by the letters X , Y or Z . To a register X and its respective Hilbert space \mathcal{H}_X , we associate an orthonormal basis $\{|x\rangle_X\}$ where $x \in \mathcal{X}$ and \mathcal{X} is a set with cardinality $|\mathcal{X}| = d_X$.

A *probability distribution* on \mathcal{X} is a map $P_X : \mathcal{X} \rightarrow [0, 1]$ such that $\sum_x P_X(x) = 1$. It can be represented as a quantum state ρ_X on X , i.e.

$$\rho_X = \sum_x P_X(x) |x\rangle\langle x|_X.$$

It follows from the properties of P_X that this state is positive semi-definite and has unit trace. In classical probability theory, a register corresponds to a discrete *random variable* and the state of a register to the probability distribution over the random variable. In this sense, quantum mechanics can be seen as a generalization of classical probability theory. Moreover, registers will become important to describe the outcomes of measurements, as we will see in the following.

2.2.2 Separated Systems

Consider two separate quantum systems, A and B , modeled by a Hilbert space \mathcal{H}_A and \mathcal{H}_B , respectively. We can prepare d_A different perfectly distinguishable states on system A and d_B different perfectly distinguishable states on system B . On the joint system AB , we can thus prepare $d_A d_B$ different perfectly distinguishable states. Hence, according to Postulate 1, the dimension of the *joint system* is $d_{AB} = d_A d_B$ and it can be modeled using the Hilbert space $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$ of dimension $d_A d_B$. *Joint states* of the two systems are then described by normalized positive semi-definite operators $\rho_{AB} \in \mathcal{S}_=(\mathcal{H}_{AB})$.

Given a state ρ_{AB} on the joint system, what are the states of the individual systems A and B ? As we have seen before, the state of the system A is fully characterized by the linear functionals acting on it, i.e. the functionals $\text{tr}(\rho_{AB}L_A) = \text{tr}(\rho_{AB}(L_A \otimes \mathbb{1}_B)) = \text{tr}(\text{tr}_B(\rho_{AB})L_A)$. Here, we introduced the notation $\text{tr}_B \equiv \text{tr}_{\mathcal{H}_B}$ for the partial trace over the subsystem B . Since the state on A is thus fully characterized by linear functionals on the operator $\text{tr}_B(\rho_{AB})$, we define the marginal state or *marginal* of ρ_{AB} on A as $\rho_A := \text{tr}_B(\rho_{AB})$. Similarly, we define the marginal of ρ_{AB} on B as $\rho_B := \text{tr}_A(\rho_{AB})$. In the following, the marginals are introduced implicitly with the joint state.¹

Entanglement

We call a state ρ_{AB} of a joint quantum system AB *separable* if it can be written in the form

$$\rho_{AB} = \sum_k \sigma_A^k \otimes \tau_B^k, \quad \text{where } \sigma_A^k \in \mathcal{P}(\mathcal{H}_A) \text{ and } \tau_B^k \in \mathcal{P}(\mathcal{H}_B). \quad (2.12)$$

Otherwise, it is *entangled*. The occurrence of entangled states is one of the most intriguing properties of the formalism of quantum mechanics.

The prime example of an entangled state is the *maximally entangled* state

$$|\gamma\rangle_{AA'} = \frac{1}{d_A} \sum_i |e_i\rangle_A \otimes |e_i\rangle_{A'},$$

where $\{|e_i\rangle\}$ is an orthonormal basis of $\mathcal{H}_A \cong \mathcal{H}_{A'}$. This state cannot be written in the form (2.12) as the following argument, due to Peres and Horodecki [Per96, HHH96], shows. Consider the super-operator $\mathcal{T} : \rho_{A'} \mapsto \rho_{A'}^T$, where the transpose is taken with regards to the basis $\{|e_i\rangle\}$ of A' . This super-operator is called the *partial transpose* (on A') and it is positive but not completely positive. Applied to separable states (2.12), it always results in positive semi-definite states $\sum_k \sigma_A^k \otimes (\tau_B^k)^T$. Applied to $\gamma_{AA'}$, however, we get

$$\mathcal{T}[\gamma_{AA'}] = \frac{1}{d_A} \sum_{i,j} |e_i\rangle\langle e_j| \otimes \mathcal{T}[|e_i\rangle\langle e_j|] = \frac{1}{d_A} \sum_{i,j} |e_i\rangle\langle e_j| \otimes |e_j\rangle\langle e_i|.$$

This state is not positive semi-definite. For example, we have

$$\langle \varphi | \mathcal{T}[\gamma_{AA'}] | \varphi \rangle = -\frac{2}{d_A}, \quad \text{where } |\varphi\rangle = |e_1\rangle \otimes |e_2\rangle - |e_2\rangle \otimes |e_1\rangle.$$

Generally, we have seen that a bipartite state is separable only if it remains positive semi-definite under the partial transpose. The converse is not true in general.

¹Whenever we introduce a state (e.g. ρ_{ABC}) on a joint system, we also implicitly introduce all its marginals (e.g. $\rho_{AB} = \text{tr}_C(\rho_{ABC})$ and $\rho_C = \text{tr}_{AB}(\rho_{ABC})$).

2. PRELIMINARIES

Classical-Quantum Systems

Joint systems where one (or more) subsystems are classical registers are of particular importance. Consider, for example, the case where a classical register X is described jointly with a quantum system A . The possible *joint states* ρ_{XA} can be written as¹

$$\rho_{XA} = \sum_x P_X(x) |x\rangle\langle x|_X \otimes \tau_A^x, \quad \text{where } \tau_A^x \in \mathcal{S}_=(\mathcal{H}_B) \quad (2.13)$$

and P_X is a probability distribution on \mathcal{X} . These states are called *classical-quantum* (CQ) states. They are of the form (2.12) and, thus, separable.

A special case occurs when the two classical registers, X and Y , are considered jointly. In this case, the states ρ_{XY} are of the form

$$\begin{aligned} \rho_{XY} &= \sum_{x,y} P_{XY}(x,y) |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \\ &= \sum_x P_X(x) |x\rangle\langle x|_X \otimes \sum_y P_Y^x(y) |y\rangle\langle y|_Y, \end{aligned}$$

where P_{XY} is a probability distribution on $\mathcal{X} \times \mathcal{Y}$ and, for each x , $P_Y^x(y) = P_{XY}(x,y)/P_X(x)$ is the conditional probability distribution on \mathcal{Y} given a fixed $x \in \mathcal{X}$. Such states allow the description of arbitrarily correlated classical random variables.

Purifications and Extensions

For any state ρ_A of a system A , we can find a *purification* on an auxiliary system A' with $\mathcal{H}_A \cong \mathcal{H}_{A'}$. A purification is a pure state $|\rho\rangle \in \mathcal{H}_{AA'}$ of the joint system AA' with the property that $\rho_A = \text{tr}_{A'}(\rho_{AA'})$. More specifically, if $\rho_A = \sum_i \lambda_i |e_i\rangle\langle e_i|_A$ is the eigenvalue decomposition of ρ_A , then a possible purification is given by $|\rho\rangle_{AA'} = \sum_i \sqrt{\lambda_i} |e_i\rangle_A \otimes |e_i\rangle_{A'}$. Purifications of ρ_A are separable if and only if ρ_A is pure. More generally, we call a (not necessarily pure) state $\rho_{AA'}$ that satisfies $\text{tr}_{A'}(\rho_{AA'}) = \rho_A$ an *extension* of the state ρ_A .

Purifications of CQ states of the form (2.13) can be constructed as

$$|\rho\rangle_{XX'AA'} = \sum_x |x\rangle_X \otimes |x\rangle_{X'} \otimes |\tau^x\rangle_{AA'},$$

where $|\tau^x\rangle$ is a purification of τ^x on AA' . We call the two registers X and X' *coherent classical*. In particular, the marginal states ρ_{XAB} and $\rho_{X'AB}$ are classical on X and X' , respectively, and they are equal with regards to the isomorphism $|x\rangle_X \mapsto |x\rangle_{X'}$.

¹This describes the full set of states that have a classical marginal on X .

2.2.3 Evolutions and Measurements

Evolution

The *evolution* of a separate quantum system is most generally described by a *quantum channel*. A quantum channel is a linear map (i.e. a super-operator) from quantum states on a system A to quantum states on a system B . Since such super-operator map quantum states onto quantum states, they must necessarily be positive and trace-preserving. Moreover, since quantum channels describe the evolution of quantum systems that may be part of a larger joint system, they are required to map positive semi-definite states of any joint system to positive semi-definite states. This implies that they are completely positive and, thus, *TP-CPMs*.

Postulate 2 (Evolution). *The evolution of quantum systems is described by trace-preserving completely positive maps.*

An important example of such a map is a time evolution. Here, system A is any system at time t_0 and system B the same system at a later time $t_1 > t_0$. In the traditional treatment of quantum mechanics, the time evolution of a system is described by a *unitary* evolution that is induced by the *Hamiltonian* of the system. A unitary evolution is a special case of a TP-CPM and describes the evolution of a closed system, i.e. a system that does not interact with any other system.

Measurement

A *measurement* of a quantum systems can be described in the above framework of general evolutions.

Postulate 3 (Measurement). *A quantum measurement is a trace-preserving completely positive map from a system to a classical register containing the measurement outcome and a system that contains the state of the system after measurement.*

Let A be a quantum system, X the classical *register* containing the measurement result and A' the system modeling the quantum system after measurement. The corresponding measurement map, $\mathcal{M} \in \mathcal{T}(\mathcal{H}_A, \mathcal{H}_{XA'})$, with $\mathcal{H}_{A'} \cong \mathcal{H}_A$, has a Kraus decomposition

$$\mathcal{M} : \rho \mapsto \sum_k E_k \rho E_k^\dagger, \quad \text{where} \quad E_k \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_{XA'}) \quad \text{and} \quad \sum_k E_k^\dagger E_k = \mathbb{1}_A.$$

Since the resulting state is required to be classical on the register X , we further know that

$$\mathcal{M}[\rho_A] = \sum_x P_X(x) |x\rangle\langle x| \otimes \tau_{A'}^x \quad \text{where} \quad \tau_{A'}^x \in \mathcal{S}_=(\mathcal{H}_{A'}). \quad (2.14)$$

2. PRELIMINARIES

Here, $P_X(x)$ is the *probability* that the outcome “ x ” is measured and $\tau_{A'}^x$ is the state of the system conditioned on the event that the outcome “ x ” has been measured. Due to (2.14), the Kraus operators E_k necessarily have the form $E_k = |x_k\rangle \otimes F_k$, where $F_k \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_{A'}) \cong \mathcal{L}(\mathcal{H}_A)$ and, thus,

$$\tau_{A'}^x = \frac{1}{P(x)} \sum_{k:x_k=x} F_k \rho_A F_k^\dagger \quad \text{and} \quad P_X(x) = \text{tr} \left(\sum_{k:x_k=x} F_k^\dagger F_k \rho \right).$$

The second equality follows from the fact that $\tau_{A'}^x$ is normalized.

Hence, a measurement is fully specified by the operators $\{F_k\}$ and the partitioning $\{k : x_k = x\}$. Note that this viewpoint is consistent; specifically, P_X is a probability distribution since

$$0 \leq \sum_{k:x_k=x} F_k^\dagger F_k \leq \mathbb{1}_A \quad \text{and} \quad \sum_x \sum_{k:x_k=x} F_k^\dagger F_k = \sum_k F_k^\dagger F_k = \mathbb{1}_A.$$

Often we are not interested in the specific state after measurement but only the probability distribution the measurement induces on the register X . In this case, the measurement is fully characterized by the operators $M_x = \sum_{k:x_k=x} F_k F_k^\dagger$, which define the probability $P_X(x) = \text{tr}(M_x \rho)$. The set $\{M_x\}$ is called a *positive operator-valued measurement* (POVM).

Definition 2.3 (POVM). *A positive operator-valued measurement on a quantum system A is a set $\{M_x\}$ with $M_x \in \mathcal{P}(\mathcal{H}_A)$ and $\sum_x M_x = \mathbb{1}_A$. The corresponding measurement TP-CPM $\mathcal{M} : \mathcal{H}_A \rightarrow \mathcal{H}_X$ is given by $\rho \mapsto \sum_x \text{tr}_{A'}(M_x \rho) |x\rangle\langle x|$ and the operators M_x are called POVM elements.*

We often use the following *Stinespring dilation* of POVM measurements. The isometry $U \in \mathcal{U}(\mathcal{H}_A, \mathcal{H}_{A'X'})$ maps the state to the classical register X containing the measurement outcome, a *coherent classical* copy of it, X' , and a possible post-measurement state on A' . The post-measurement states that result from a POVM are not unique; however, conventionally, one takes $x_k = k$ where k is chosen from the same set as x and $F_k = \sqrt{M_x}$. Thus,

$$\mathcal{M}[\rho] = \text{tr}_{X'A'}(U \rho U^\dagger) \quad \text{where} \quad U = \sum_x |x\rangle_X \otimes |x\rangle_{X'} \otimes \sqrt{M_x}.$$

A special case of a POVM occurs when the *POVM elements* are projectors, i.e. $M_x = M_x^2$. We call such a measurement a *projective measurement*. If these projections are of rank 1, e.g. $M_x = |x\rangle\langle x|_X$, the corresponding measurement TP-CPM takes on the simple form $\mathcal{M} : \rho_A \mapsto \sum_x \langle x | \rho_A | x \rangle |x\rangle\langle x|_X$.

Comparing this to the traditional treatment of quantum measurements, we note that it is not necessary to introduce a separate formalism for measurements and that we can treat measurements as a special case of an evolution. This is possible since we always consider the outcome of a measurement as a *random variable* that is correlated with the state of the system after measurement and do not condition the resulting quantum state on a particular measurement outcome.

2.3 Mathematical Toolkit

This section covers the most important mathematical tools used throughout this thesis. The (smooth) min- and max-entropy introduced in Chapter 4 can be formulated as semi-definite programs. Operator monotone functions are used to explore properties of another class of entropies, generalizations of Rényi entropies, in Chapter 6.

2.3.1 Semi-Definite Programs

This overview is based on the Watrous Lecture Notes [Wat08] and all proofs can be found there. A *semi-definite program* (SDP) is a triple $\{A, B, \Psi\}$, where $A \in \mathcal{H}(\mathcal{H})$, $B \in \mathcal{H}(\mathcal{H}')$ and $\Psi \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ is a super-operator from \mathcal{H} to \mathcal{H}' that preserves Hermiticity. The following two optimization problems are associated with the semi-definite program.

primal problem	dual problem
$\text{minimize : } \langle A, X \rangle$ $\text{subject to : } \Psi[X] \geq B$ $X \in \mathcal{P}(\mathcal{H})$	$\text{maximize : } \langle B, Y \rangle$ $\text{subject to : } \Psi^\dagger[Y] \leq A$ $Y \in \mathcal{P}(\mathcal{H}')$

We call an operator $X \in \mathcal{P}(\mathcal{H})$ primal feasible if it satisfies $\Psi[X] \geq B$. Similarly, we say that $Y \in \mathcal{P}(\mathcal{H}')$ is dual feasible if $\Psi^\dagger[Y] \leq A$. Moreover, we denote the optimal solution of the primal problem with α and the optimal solution of the dual problem with β . Formally,

$$\begin{aligned} \alpha &:= \inf \{ \langle A, X \rangle : X \in \mathcal{P}(\mathcal{H}), \Psi[X] \geq B \} \\ \beta &:= \sup \{ \langle B, Y \rangle : Y \in \mathcal{P}(\mathcal{H}'), \Psi^\dagger[Y] \leq A \}. \end{aligned} \quad (2.15)$$

The following two theorems provide a relation between the primal and dual problem of an SDP.

Theorem 2.5 (Weak Duality). *Let $\{A, B, \Psi\}$ be a SDP and α, β defined as in (2.15). Then, $\alpha \geq \beta$.*

This implies that every dual feasible operator Y provides a lower bound of $\langle B, Y \rangle$ on α and every primal feasible operator X provides an upper bound of $\langle A, X \rangle$ on β .

Theorem 2.6 (Strong Duality). *Let $\{A, B, \Psi\}$ be a SDP and α, β defined as in (2.15). Then the following holds:*

- *If α is finite and there exists an operator $Y > 0$ such that $\Psi^\dagger[Y] < A$, then $\alpha = \beta$ and there exists a primal feasible X such that $\langle A, X \rangle = \alpha$.*
- *If β is finite and there exists an operator $X > 0$ such that $\Psi[X] > B$, then $\alpha = \beta$ and there exists a dual feasible Y such that $\langle B, Y \rangle = \beta$.*

2. PRELIMINARIES

Optimization problems that can be formulated as semi-definite programs can be efficiently solved numerically.¹

2.3.2 Operator Monotone Functions

Here, we discuss some useful properties of operator *monotone*, *concave* and *convex* functions. This section is largely based on Chapter V of Bhatia [Bha97] and we simply repeat the results here in the form we need in later chapters.

Operator monotone functions preserve the partial order on operators induced by ‘ \geq ’ and are, thus, necessarily monotone.

Definition 2.4. Let $\Omega \subseteq \mathbb{R}$. A function $f : \Omega \rightarrow \mathbb{R}$ is operator monotone on Ω if $A \geq B$ implies $f(A) \geq f(B)$ for any Hermitian operators A, B with eigenvalues in Ω .

Similarly, operator concave and convex functions generalize the concept of concavity and convexity to operators.

Definition 2.5. Let Ω be an interval on \mathbb{R} . A function $f : \Omega \rightarrow \mathbb{R}$ is operator concave on Ω if $f(\mu A + (1 - \mu)B) \geq \mu f(A) + (1 - \mu)f(B)$ for all Hermitian operators A, B with eigenvalues in Ω and all $\mu \in [0, 1]$. The function f is operator convex on Ω if $-f$ is operator concave on Ω .

Prominent examples of such functions include (cf. Chapter V in [Bha97])

- The logarithm function, which is operator monotone on \mathbb{R}^+ .
- The function $h : t \mapsto -t \log t$ with its extension to $h(0) = \lim_{t \rightarrow 0} h(t) = 0$, which is operator concave on \mathbb{R}_0^+ .
- The family of functions $g_\alpha : t \mapsto t^\alpha$. These functions are operator concave and operator monotone on \mathbb{R}_0^+ for $\alpha \in (0, 1]$ and operator convex on \mathbb{R}_0^+ for $\alpha \in [1, 2]$. (Note that these functions are convex but not operator convex if $\alpha > 2$.)

We start with a straightforward application of Jensen’s inequality:

Lemma 2.7. Let Ω be an interval on \mathbb{R} and let $f : \Omega \rightarrow \mathbb{R}$ be concave on Ω . Then, for any $|\varphi\rangle \in \mathcal{H}$ with $\|\varphi\| = 1$ and $A \in \mathcal{H}(\mathcal{H})$ with eigenvalues in Ω , we have $\langle \varphi | f(A) | \varphi \rangle \leq f(\langle \varphi | A | \varphi \rangle)$.

Proof. Using the eigenvalue decomposition $A = \sum_i \lambda_i |e_i\rangle\langle e_i|$, we get

$$\langle \varphi | f(A) | \varphi \rangle = \sum_i f(\lambda_i) |\langle \varphi | e_i \rangle|^2 \leq f\left(\sum_i \lambda_i |\langle \varphi | e_i \rangle|^2\right) = f(\langle \varphi | A | \varphi \rangle),$$

where we used that $\sum_i |\langle \varphi | e_i \rangle|^2 = 1$. □

¹For example, using the SeDuMi solver [SED] and YALMIP front-end [Löf04].

A comprehensive generalization of Jensen's inequality to operator convex functions is given in [HP03]. We state a specialized version here for completeness.

Lemma 2.8 (Operator Jensen's Inequality). *Let Ω be an interval on \mathbb{R} and let $f : \Omega \rightarrow \mathbb{R}$ be continuous and operator concave on Ω . Then, for any isometry $U : \mathcal{H} \rightarrow \mathcal{H}'$ and $A \in \mathcal{H}(\mathcal{H})$ with eigenvalues in Ω , we have*

$$Uf(A)U^\dagger \leq f(UAU^\dagger).$$

2. PRELIMINARIES

Chapter 3

The Purified Distance

This chapter is based on [TCR10], where the purified distance was first proposed as a metric on the space of sub-normalized quantum states. The usefulness of the purified distance will become apparent when it is applied to define the smooth min- and max-entropies in Chapter 4, providing them with natural properties such as invariance under local isometries and various data processing inequalities.

3.1 Introduction and Related Work

Smooth *entropies*, evaluated for a quantum state ρ , are defined indirectly via an optimization (either a maximization or a minimization) of an underlying unsmoothed entropy over a set of states that are ε -close to ρ , where ε is a small *smoothing parameter*. The resulting quantities are called ε -*smooth* entropies. (See, for example, Chapter 4, where the ε -smooth min- and max-entropies are defined in this way.)

Consequently, various definitions of such sets of close states—subsequently called ε -*balls*—have appeared in the literature. However, to the best of our knowledge, none of the existing definitions simultaneously exhibit the following two properties that are of particular importance in the context of smooth entropies in the quantum regime.

- Firstly, the definition of the ε -smooth entropies should be independent of the Hilbert spaces used to represent the quantum state ρ .

In particular, embedding ρ into a larger Hilbert space prior to smoothing should leave the ε -smooth entropies unchanged. Note that, in general, embedding ρ into a larger Hilbert space offers more flexibility for smoothing as more dimensions orthogonal to the support of ρ become available for the optimization. Indeed, for some ε -balls that contain only normalized states, smoothing outside the support

3. THE PURIFIED DISTANCE

becomes advantageous and the smooth entropy thus depends on the Hilbert space representation of ρ .

We can avoid this problem by including sub-normalized quantum states in the ε -balls.

- Secondly, it will be important that we can define a ball of pure states that contains purifications of all the states in a given ε -ball. This will allow us to establish duality relations between smooth entropies and is achieved by using a fidelity-based metric to determine ε -closeness.

3.1.1 Main Contributions

The following sections introduce a new metric on sub-normalized quantum states, the *purified distance*. We call two quantum states ρ and τ ε -close, denoted $\rho \approx_\varepsilon \tau$, if and only if the purified distance between them is at most ε . The purified distance has various interesting properties, among them are the following.

Result 1 (Purified Distance). *The purified distance is a metric on sub-normalized states and has the following properties:*

- If at least one of the states is normalized, it can be expressed in terms of the fidelity as $P(\rho, \tau) = \sqrt{1 - F(\rho, \tau)^2}$.
- It is an upper bound to the trace distance.
- For any trace non-increasing CPM \mathcal{E} and any states ρ and σ , we have $\rho \approx_\varepsilon \tau \implies \mathcal{E}[\rho] \approx_\varepsilon \mathcal{E}[\tau]$.
- If ρ_{AB} is a state and $\sigma_A \approx_\varepsilon \rho_A$ is close to its marginal, then there always exists an extension σ_{AB} of σ_A with $\sigma_{AB} \approx_\varepsilon \rho_{AB}$.

3.1.2 Outline

Section 3.2 introduces two metrics on the set of sub-normalized states, the generalized trace distance and the purified distance. In Section 3.3, we discuss various properties of the purified distance, including its relation to the generalized trace distance, its monotonicity under trace non-increasing CPMs and an adaption of Uhlmann's theorem to the purified distance. Section 3.4 then explains some notational conventions, made possible by the use of the purified distance as a metric, that should help the reader through the remaining chapters.

3.2 Two Metrics for Quantum States

For the remainder of this chapter, let \mathcal{H} be an arbitrary finite-dimensional Hilbert space.¹ The two most common measures of distance between normalized quantum states are the *trace distance* and the *fidelity*.

3.2.1 Generalized Trace Distance

We start by introducing a straight-forward generalization of the trace distance to sub-normalized quantum states.

Definition 3.1. For $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$, we define the generalized trace distance between ρ and τ as

$$D(\rho, \tau) := \max \{ \operatorname{tr}\{\rho - \tau\}_+, \operatorname{tr}\{\tau - \rho\}_+ \}.$$

The generalized trace distance can be expressed alternatively in terms of the Schatten 1-norm as

$$D(\rho, \tau) = \frac{1}{2} \|\rho - \tau\|_1 + \frac{1}{2} |\operatorname{tr} \rho - \operatorname{tr} \tau|$$

and it is easy to verify that it is a *metric* on $\mathcal{L}(\mathcal{H})$. In the case where both ρ and τ are normalized states, we recover the usual definition of the trace distance, $D(\rho, \tau) := \frac{1}{2} \|\rho - \tau\|_1$. Furthermore, the trace distance has a physical interpretation as the *distinguishing advantage* between the two states. In other words, the probability $p_{\text{dist}}(\rho, \tau)$ of correctly distinguishing between two equiprobable states ρ and τ by any measurement is upper bounded by [NC00]

$$p_{\text{dist}}(\rho, \tau) \leq \frac{1}{2} (1 + D(\rho, \tau)). \quad (3.1)$$

3.2.2 Generalized Fidelity

On the other hand, various metrics are derived from the *fidelity*, which is given as $F(\rho, \tau) = \|\sqrt{\rho}\sqrt{\tau}\|_1$ for two normalized states ρ and τ . We will not use the letter F to denote the fidelity hereafter, instead reserving it for the generalized fidelity defined below. (We will also see that the two quantities agree if at least one state is normalized.) The fidelity has many interesting properties, some of which we will list here for further reference. The most important properties of the fidelity are summarized in Table 3.1.

Here, we propose a generalization of the fidelity to sub-normalized states. The generalization is motivated by the observation that sub-normalized states can be thought of as normalized states on a larger space

¹Note also that most results of this chapter have recently been generalized to the framework of general von Neumann algebras [BFS11].

3. THE PURIFIED DISTANCE

The following holds for arbitrary positive operators $\rho, \tau, \sigma \in \mathcal{P}(\mathcal{H})$.

- i. The fidelity is *symmetric* in its arguments, $\|\sqrt{\rho}\sqrt{\tau}\|_1 = \|\sqrt{\tau}\sqrt{\rho}\|_1$.
- ii. The fidelity is monotonously increasing under the application of TP-CPMs (cf. e.g. [NC00], Theorem 9.6). This implies that, for any TP-CPM \mathcal{E} , we have

$$\sqrt{\text{tr } \rho} \sqrt{\text{tr } \tau} \geq \left\| \sqrt{\mathcal{E}[\rho]} \sqrt{\mathcal{E}[\tau]} \right\|_1 \geq \|\sqrt{\rho}\sqrt{\tau}\|_1. \quad (3.2)$$

To get the first inequality, we used that tr is a TP-CPM.

- iii. *Uhlmann's theorem* [Uhl85] states that, for any purification φ of ρ ,

$$\|\sqrt{\rho}\sqrt{\tau}\|_1 = \max_{|\vartheta\rangle} |\langle \varphi | \vartheta \rangle|, \quad (3.3)$$

where the maximum is taken over all *purifications* ϑ of τ .

- iv. For any *projector* $\Pi \in \mathcal{P}(\mathcal{H})$, we have

$$\|\sqrt{\Pi\rho\Pi}\sqrt{\tau}\|_1 = \|\sqrt{\Pi\rho\Pi}\sqrt{\Pi\tau\Pi}\|_1 = \|\sqrt{\rho}\sqrt{\Pi\tau\Pi}\|_1. \quad (3.4)$$

- v. For any $\sigma \geq \rho$, we have $\|\sqrt{\sigma}\sqrt{\tau}\|_1 \geq \|\sqrt{\rho}\sqrt{\tau}\|_1$.
- vi. For states $\rho = \rho_1 \oplus \rho_2$ and $\tau = \tau_1 \oplus \tau_2$, where $\rho_1, \tau_1 \in \mathcal{P}(\mathcal{H}_1)$ and $\rho_2, \tau_2 \in \mathcal{P}(\mathcal{H}_2)$, we have

$$\|\sqrt{\rho}\sqrt{\tau}\|_1 = \|\sqrt{\rho_1}\sqrt{\tau_1}\|_1 + \|\sqrt{\rho_2}\sqrt{\tau_2}\|_1.$$

Table 3.1: *Properties of the Fidelity.*

which are projected onto a subspace. Hence, we define the generalized fidelity as the supremum of the fidelity between such normalized states.

Definition 3.2. For $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$, we define the generalized fidelity between ρ and τ as

$$F(\rho, \tau) := \sup_{\mathcal{H}'} \sup_{\bar{\rho}, \bar{\tau} \in \mathcal{S}_{=}(\mathcal{H}')} \left\| \sqrt{\bar{\rho}} \sqrt{\bar{\tau}} \right\|_1, \quad (3.5)$$

where the supremum is taken over all embeddings V of \mathcal{H} into \mathcal{H}' and all states $\bar{\rho}, \bar{\tau} \in \mathcal{S}_{=}(\mathcal{H}')$, such that ρ and τ are images of $\bar{\rho}$ and $\bar{\tau}$ under V^\dagger . (Namely, the states satisfy $V^\dagger \bar{\rho} V = \rho$ and $V^\dagger \bar{\tau} V = \tau$.)

This expression reduces to the fidelity when at least one state is normalized. To see this, consider the following alternative expression for the generalized fidelity.

Lemma 3.1. *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$. Then,*

$$F(\rho, \tau) = F(\hat{\rho}, \hat{\tau}) = \|\sqrt{\rho}\sqrt{\tau}\|_1 + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \tau)}, \quad (3.6)$$

where $\hat{\rho} := \rho \oplus (1 - \text{tr } \rho)$ and $\hat{\tau} := \tau \oplus (1 - \text{tr } \tau)$.

Proof. Let $U : \mathcal{H} \rightarrow \mathcal{H}'$, $\bar{\rho}$ and $\bar{\tau}$ be an arbitrary candidate for the supremum in (3.5). Moreover, let \mathcal{E} be the pinching $\rho \mapsto \Pi\rho\Pi + \Pi^\perp\rho\Pi^\perp$, where $\Pi := UU^\dagger$ projects onto the image of U and $\Pi^\perp := \mathbb{1} - \Pi$ is its orthogonal complement on \mathcal{H}' . Then, due to the monotonicity property (3.2), we find

$$\begin{aligned} \|\sqrt{\bar{\rho}}\sqrt{\bar{\tau}}\|_1 &\leq \|\sqrt{\mathcal{E}[\bar{\rho}]}\sqrt{\mathcal{E}[\bar{\tau}]}\|_1 \\ &= \|\sqrt{\Pi\bar{\rho}\Pi}\sqrt{\Pi\bar{\tau}\Pi}\|_1 + \|\sqrt{\Pi^\perp\bar{\rho}\Pi^\perp}\sqrt{\Pi^\perp\bar{\tau}\Pi^\perp}\|_1 \\ &\leq \|\sqrt{\rho}\sqrt{\tau}\|_1 + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \tau)}. \end{aligned}$$

In particular, the r.h.s. is an upper bound on $F(\rho, \tau)$. Finally, it is easy to verify that this upper bound is achieved with the choice $\mathcal{H}' \cong \mathcal{H} \oplus \mathbb{C}$ as well as $\hat{\rho}$ and $\hat{\tau}$. \square

3.2.3 Purified Distance

Next, we define a distance measure based on the fidelity, analogously to the one proposed in [GLN04, Ras06]¹.

Definition 3.3 (Purified Distance). *For $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$, we define the purified distance between ρ and τ as*

$$P(\rho, \tau) := \sqrt{1 - F(\rho, \tau)^2}.$$

Other distance measures based on the fidelity have been investigated in the literature. In particular, the *Bures metric* [Bur69], $B(\rho, \tau)^2 := 2(1 - F(\rho, \tau))$, and the *angular distance*, $A(\rho, \tau) := \arccos F(\rho, \tau)$ [NC00]. We prefer the purified distance because it constitutes an upper bound on the trace distance as we will see below and thus inherits its operational interpretation as an upper bound on the *distinguishing advantage*.

The name ‘‘Purified Distance’’ is motivated by the fact that, for normalized states $\rho, \tau \in \mathcal{S}_{=}(\mathcal{H})$, we can write $P(\rho, \tau)$ as the minimum trace distance between purifications $|\varphi\rangle$ of ρ and $|\vartheta\rangle$ of τ . More precisely, using Uhlmann’s theorem (3.3), we have

$$\begin{aligned} P(\rho, \tau) &= \sqrt{1 - F(\rho, \tau)^2} = \sqrt{1 - \max_{\varphi, \vartheta} |\langle \varphi | \vartheta \rangle|^2} \\ &= \min_{\varphi, \vartheta} \sqrt{1 - |\langle \varphi | \vartheta \rangle|^2} = \min_{\varphi, \vartheta} D(\varphi, \vartheta). \end{aligned}$$

¹The quantity $C(\rho, \tau) = \sqrt{1 - F^2(\rho, \tau)}$ is introduced in [GLN04], where the authors also show that it is a metric on $\mathcal{S}_{=}(\mathcal{H})$. In [Ras06], the same quantity is called sine distance and some of its properties are explored.

3. THE PURIFIED DISTANCE

The purified distance is a *metric* on the set of sub-normalized states according to Definition 2.1.

Proposition 3.2. *The purified distance $P(\cdot, \cdot)$ is a metric on $\mathcal{S}_{\leq}(\mathcal{H})$.*

Proof. Let $\rho, \tau, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$. The condition $P(\rho, \tau) = 0 \iff \rho = \tau$ can be verified by inspection, and symmetry $P(\rho, \tau) = P(\tau, \rho)$ follows from the symmetry of the fidelity.

It remains to show the triangle inequality, $P(\rho, \tau) \leq P(\rho, \sigma) + P(\sigma, \tau)$. Using Lemma 3.1, the generalized fidelities between ρ , τ and σ can be expressed as fidelities between the corresponding extensions $\hat{\rho}$, $\hat{\tau}$ and $\hat{\sigma}$. We employ the triangle inequality of the *angular distance*, which can be expressed in terms of the purified distance as $A(\rho, \tau) = \arccos F(\rho, \tau) = \arcsin P(\rho, \tau)$.¹ This leads to

$$\begin{aligned} P(\rho, \tau) &= \sin A(\hat{\rho}, \hat{\tau}) \\ &\leq \sin(A(\hat{\rho}, \hat{\sigma}) + A(\hat{\sigma}, \hat{\tau})) \\ &= \sin A(\hat{\rho}, \hat{\sigma}) \cos A(\hat{\sigma}, \hat{\tau}) + \sin A(\hat{\sigma}, \hat{\tau}) \cos A(\hat{\rho}, \hat{\sigma}) \end{aligned} \quad (3.7)$$

$$\begin{aligned} &= P(\rho, \sigma)F(\sigma, \tau) + P(\sigma, \tau)F(\rho, \sigma) \\ &\leq P(\rho, \sigma) + P(\sigma, \tau), \end{aligned} \quad (3.8)$$

where we employed the trigonometric addition formula to get (3.7). \square

Note that the purified distance is not an intrinsic metric, i.e. given two states ρ, τ with $P(\rho, \tau) \leq \varepsilon$ it is in general not possible to find intermediate states σ^λ with $P(\rho, \sigma^\lambda) = \lambda\varepsilon$ and $P(\sigma^\lambda, \tau) = (1 - \lambda)\varepsilon$. In this sense, the above triangle inequality is not tight. It is thus sometimes useful to employ Eq. (3.8) instead. For example, given three states $\rho, \tau, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ and $0 \leq \varepsilon, \bar{\varepsilon} \leq 1$, we find that $P(\rho, \sigma) \leq \varepsilon$ and $P(\sigma, \tau) \leq \bar{\varepsilon}$ implies

$$P(\rho, \tau) \leq \varepsilon\sqrt{1 - \bar{\varepsilon}^2} + \bar{\varepsilon}\sqrt{1 - \varepsilon^2} \quad (3.9)$$

if $\arcsin \varepsilon + \arcsin \bar{\varepsilon} \leq \frac{\pi}{2}$. This bound is plotted in Figure 3.1.

3.3 Properties of the Purified Distance

The purified distance allows simple upper and lower bounds in terms of the generalized trace distance.

Proposition 3.3. *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$. Then*

$$D(\rho, \tau) \leq P(\rho, \tau) \leq \sqrt{2D(\rho, \tau) - D(\rho, \tau)^2} \leq \sqrt{2D(\rho, \tau)}.$$

¹A proof that A is a metric for normalized states is given in [NC00].

3.3 Properties of the Purified Distance

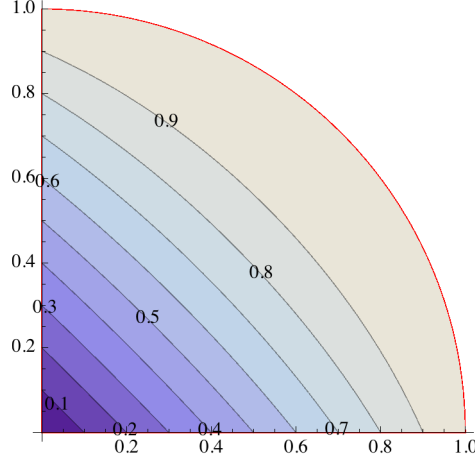


Figure 3.1: *Improved Triangle Inequality*. Contour plot of the expression in Eq. (3.9) with axes ε and $\bar{\varepsilon}$. The boundary contains all pairs $\{\varepsilon, \bar{\varepsilon}\}$ that lead to the trivial bound $P(\rho, \tau) \leq 1$. This shows that Eq. (3.8) gives significantly tighter results than the triangle inequality for the purified distance if ε and $\bar{\varepsilon}$ get large. For example, if $P(\rho, \sigma) = P(\sigma, \tau) = 0.5$, we find $P(\rho, \tau) \leq 0.87$ instead of the trivial $P(\rho, \tau) \leq 1$.

Proof. We express the quantities using the normalized extensions $\hat{\rho}$ and $\hat{\tau}$ of Lemma 3.1 to get

$$\begin{aligned} P(\rho, \tau) &= \sqrt{1 - F(\hat{\rho}, \hat{\tau})^2} \geq D(\hat{\rho}, \hat{\tau}) = D(\rho, \tau) \quad \text{and} \\ P(\rho, \tau)^2 &= 1 - F(\hat{\rho}, \hat{\tau})^2 \leq 1 - (1 - D(\hat{\rho}, \hat{\tau}))^2 \\ &= 2D(\rho, \tau) - D(\rho, \tau)^2 \leq 2D(\rho, \tau), \end{aligned}$$

where we employed the Fuchs – van de Graaf inequalities $1 - F(\hat{\rho}, \hat{\tau}) \leq D(\hat{\rho}, \hat{\tau}) \leq \sqrt{1 - F(\hat{\rho}, \hat{\tau})^2}$ (cf. [Fuc96] and [NC00], Section 9.2.3). \square

One very useful property of the purified distance is that it does not increase under simultaneous application of a quantum operation on both states. We consider the class of *trace non-increasing* CPMs, which includes *TP-CPMs*, *projections* and *partial isometries*.

Theorem 3.4 (Monotonicity of Purified Distance). *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\mathcal{E} : \mathcal{H} \rightarrow \mathcal{H}'$ a trace non-increasing CPM. Then,*

$$F(\mathcal{E}[\rho], \mathcal{E}[\tau]) \geq F(\rho, \tau) \quad \text{and} \quad P(\mathcal{E}[\rho], \mathcal{E}[\tau]) \leq P(\rho, \tau).$$

Proof. Remember that a trace non-increasing CPM $\mathcal{E} : \mathcal{H} \rightarrow \mathcal{H}'$ can be decomposed into an isometry $U : \mathcal{H} \rightarrow \mathcal{H}' \otimes \mathcal{H}''$ followed by a projection $\Pi \in \mathcal{P}(\mathcal{H}' \otimes \mathcal{H}'')$ and a partial trace over \mathcal{H}'' (cf. Lemma 2.4). Isometries

3. THE PURIFIED DISTANCE

and the partial trace are TP-CPMs and, hence, it suffices to show that $F(\rho, \tau) \leq F(\mathcal{E}[\rho], \mathcal{E}[\tau])$ for TP-CPMs and projections.

First, let \mathcal{E} be a TP-CPM. Using (3.6) and the monotonicity under TP-CPMs of the fidelity (3.2), we see that

$$\begin{aligned} F(\rho, \tau) &= \|\sqrt{\rho}\sqrt{\tau}\|_1 + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \tau)} \\ &\leq \|\sqrt{\mathcal{E}[\rho]}\sqrt{\mathcal{E}[\tau]}\|_1 + \sqrt{(1 - \text{tr } \rho)(1 - \text{tr } \tau)} \\ &= F(\mathcal{E}[\rho], \mathcal{E}[\tau]). \end{aligned}$$

Next, consider a projection $\Pi \in \mathcal{P}(\mathcal{H})$ and the CPM $\mathcal{E} : \rho \mapsto \Pi\rho\Pi$. Following the definition of the generalized fidelity in (3.5), we write $F(\rho, \tau) = \sup \|\sqrt{\bar{\rho}}\sqrt{\bar{\tau}}\|_1$, where the supremum is taken over all normalized extensions $\{\mathcal{H}', \bar{\rho}, \bar{\tau}\}$ of $\{\mathcal{H}, \rho, \tau\}$. Since all normalized extensions of $\{\mathcal{H}, \rho, \tau\}$ are obviously also normalized extensions of $\{\text{supp } \{\Pi\}, \Pi\rho\Pi, \Pi\tau\Pi\}$, we find $F(\Pi\rho\Pi, \Pi\tau\Pi) \geq F(\rho, \tau)$.

Finally, the second statement trivially follows from the first one by definition of the purified distance. \square

The main advantage of the purified distance (and other metrics based on the fidelity) over the trace distance is that we can always find extensions and purifications without increasing the distance. This is captured in the following two results.

Theorem 3.5 (Uhlmann's Theorem for Purified Distance). *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$, $\mathcal{H}' \cong \mathcal{H}$ and $|\varphi\rangle \in \mathcal{H} \otimes \mathcal{H}'$ be a purification of ρ . Then, there exists a purification $|\vartheta\rangle \in \mathcal{H} \otimes \mathcal{H}'$ of τ that satisfies $P(\rho, \tau) = P(\varphi, \vartheta)$.*

Proof. We use Uhlmann's theorem for the fidelity (3.3) to choose $|\vartheta\rangle \in \mathcal{H} \otimes \mathcal{H}'$ such that $\|\sqrt{\rho}\sqrt{\tau}\|_1 = |\langle\varphi|\vartheta\rangle|$ holds. Then, due to (3.6), we have $F(\rho, \tau) = F(\varphi, \vartheta)$ as well as $P(\rho, \tau) = P(\varphi, \vartheta)$. \square

Corollary 3.6. *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ be an extension of ρ . Then, there exists an extension $\bar{\tau} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ of τ with $P(\rho, \tau) = P(\bar{\rho}, \bar{\tau})$.*

Proof. Let $\mathcal{H}'' \cong \mathcal{H} \otimes \mathcal{H}'$ be an auxiliary Hilbert space and $\varphi \in \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{H}''$ be a purification of $\bar{\rho}$. We introduce a purification $\vartheta \in \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{H}''$ of τ with $P(\varphi, \vartheta) = P(\rho, \tau)$ using Uhlmann's theorem for the purified distance above and define $\bar{\tau} = \text{tr}_{\mathcal{H}''}(\vartheta)$. However, due to monotonicity (cf. Theorem 3.4), we have $P(\varphi, \vartheta) \geq P(\bar{\rho}, \bar{\tau}) \geq P(\rho, \tau)$, which implies that all three quantities must be equal. \square

The next lemma offers an explicit construction that satisfies Corollary 3.6. This was shown in [DBWR10] and we provide the proof here for completeness.

Lemma 3.7. *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ an extensions of ρ . Then, there exists an operator $X \in \mathcal{L}(\mathcal{H})$ such that $\bar{\tau} = X\bar{\rho}X^\dagger \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ is an extension of τ with $P(\rho, \tau) = P(\bar{\rho}, \bar{\tau})$.*

Proof. We show the theorem for the case when $\bar{\rho}$ is pure. The general statement then follows from the same arguments outlined in the proof of Corollary 3.6. Let $U|\sqrt{\rho}\sqrt{\tau}|$ be the polar decomposition of $\sqrt{\rho}\sqrt{\tau}$. Then, we define $X := \tau^{1/2}U\rho^{-1/2}$ using the generalized inverse. Clearly, $\bar{\tau} = X\bar{\rho}X^\dagger$ is an extension of τ as $\text{tr}_{\mathcal{H}'}(\bar{\tau}) = X\rho X^\dagger = \tau$. Furthermore, we find

$$\|\sqrt{\bar{\rho}}\sqrt{\bar{\tau}}\|_1 = |\langle \bar{\rho} | \bar{\tau} \rangle| = |\langle \bar{\rho} | X | \bar{\rho} \rangle| = |\text{tr}(X\rho)| = |\text{tr}(U\sqrt{\rho}\sqrt{\tau})| = \|\sqrt{\rho}\sqrt{\tau}\|_1.$$

The equality of the purified distance then follows by Lemma 3.1 and the definition of the purified distance. \square

The following lemma (see also [BCC⁺10, TR11]) gives a bound on the distance between a state and the projection of that state onto a subspace.

Lemma 3.8. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\Pi \in \mathcal{P}(\mathcal{H})$ be a projector, then*

$$P(\rho, \Pi\rho\Pi) \leq \sqrt{2 \text{tr}(\Pi^\perp \rho) - \text{tr}(\Pi^\perp \rho)^2},$$

where $\Pi^\perp := \mathbb{1} - \Pi$ is the complement of Π on \mathcal{H} .

Proof. The generalized fidelity between the two states can be bounded using $\text{tr}(\Pi\rho) \leq \text{tr}(\rho)$ and $\|\sqrt{\rho}\sqrt{\Pi\rho\Pi}\|_1 = \text{tr}(\Pi\rho)$, which follows from (3.4). We have

$$F(\rho, \Pi\rho\Pi) \geq \text{tr}(\Pi\rho) + 1 - \text{tr}(\rho) = 1 - \text{tr}(\Pi^\perp \rho).$$

The desired bound on the purified distance follows from its definition. \square

3.4 Notational Conventions

We will often use the notation $\rho \approx_\varepsilon \tau$ (in words, “ ρ is ε -close to τ ”) to say that $P(\rho, \tau) \leq \varepsilon$, where $\varepsilon \ll 1$ is a small parameter and $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H})$.

The following intuitive notational conventions will be used widely in the following chapters and shortens many proofs. Let \mathcal{H}_{AB} be a bipartite Hilbert space. If $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\tilde{\rho}_A \in \mathcal{S}_{\leq}(\mathcal{H}_A)$ are defined, we implicitly define $\tilde{\rho}_{AB}$ as the extension of $\tilde{\rho}_A$ that has minimum purified distance to ρ_{AB} according to Corollary 3.6 and Lemma 3.7. This implies that $P(\tilde{\rho}_{AB}, \rho_{AB}) = P(\tilde{\rho}_A, \rho_B)$.

Generally, states with the same greek letter will be treated in this way, where the distance of the modified letter states (e.g. $\tilde{\rho}, \hat{\sigma}, \bar{\tau}$) is always measured with regards to the unmodified letter state (e.g. ρ, σ, τ). Consistent with that, if an extension is introduced that is not of minimum distance,

3. THE PURIFIED DISTANCE

we will always use another greek letter to denote it. Due to the arguments above, it is usually sufficient to write $P(\tilde{\rho}, \rho)$ without mentioning the subspaces on which the states are compared.

These minimum distance extensions often inherit properties of the original state. For the example discussed above, if ρ is classical on B then $\tilde{\rho}$ is also classical on B . This follows directly from the monotonicity (cf. Theorem 3.4) of the purified distance under trace non-increasing maps, since $P(\mathcal{M}[\tilde{\rho}], \rho) = P(\mathcal{M}[\tilde{\rho}], \mathcal{M}[\rho]) \leq P(\tilde{\rho}, \rho)$, where \mathcal{M} measures in the classical basis of B and thus leaves ρ invariant. Using the same argument, one can show that the minimum distance extension $\tilde{\rho}_{AB}$ lives in the subspace $\mathcal{H}_A \otimes \text{supp}\{\rho_B\}$ of \mathcal{H}_{AB} .

Furthermore, if $\mathcal{E} : \mathcal{H}_A \rightarrow \mathcal{H}_B$ is a trace non-increasing map and $\tau_B = \mathcal{E}[\rho_A]$, then, the definition of a state $\tilde{\rho} \approx_\varepsilon \rho$ also implicitly defines a state $\tilde{\tau} = \mathcal{E}[\tilde{\rho}_A] \approx_\varepsilon \tau$. This will often be used when \mathcal{E} is an isometry. In this case, we often give the states the same letter, for example if $U = \sum_x |x\rangle_X \otimes |x\rangle_{X'} \langle x|_A$ is the isometry that purifies a projective measurement of the system A in the basis $\{|x\rangle\}$, we use $\rho_{XX'} = U\rho_A U^\dagger$ to denote the mapped state.

Chapter 4

Min- and Max-Entropies

This chapter formally introduces the min- and max-entropies for quantum states and discusses some of their properties. We provide a plethora of different expressions for the min- and max-entropy and introduce the interpretation of the min-entropy as a guessing probability. Moreover, we explore the classical limits of the min- and max-entropy and investigate their continuity and the concavity of the max-entropy.

4.1 Introduction and Related Work

We have seen in the introduction that generalizations of the classical *Rényi α -entropies* [Rén61] can be used to characterize different information theoretic tasks in the *one-shot* setting. For a discrete probability distribution P_X over a set \mathcal{X} , the Rényi α -entropies are defined as

$$H_\alpha(X)_P := \frac{1}{1-\alpha} \log \sum_x P_X(x)^\alpha, \quad \text{where } \alpha \in (0, 1) \cup (1, \infty).$$

These entropies have a trivial generalization to the quantum setting, which, for a state $\rho \in \mathcal{S}_=(\mathcal{H}_A)$, is given as

$$H_\alpha(A)_\rho := \frac{1}{1-\alpha} \log \text{tr}(\rho_A^\alpha). \quad (4.1)$$

The range of allowed α can be extended to include 0 and ∞ by taking the respective limits of (4.1). This leads to the expressions

$$H_\infty(A)_\rho := -\log \|A\|_\infty \quad \text{and} \quad H_0(A)_\rho := \log \text{rank}\{\rho_A\}.$$

Furthermore, taking the limit to $\alpha = 1$ from both sides reveals the *von Neumann entropy*; hence, we set $H_1 \equiv H$ and have now defined a spectrum of entropies for $\alpha \in [0, \infty]$. These entropies are monotonically decreasing in the parameter α , i.e.

$$\alpha \geq \beta \implies H_\alpha(A)_\rho \leq H_\beta(A)_\rho \quad \text{for all } \rho \in \mathcal{S}_=(\mathcal{H}_A).$$

4. MIN- AND MAX-ENTROPIES

One of the first questions to answer is now whether we need to consider the whole spectrum of Rényi α -entropies for our framework for non-asymptotic information theory. This was answered to the negative by Renner and Wolf [RW04, RW05]. They show that if we allow a small variation of the state of the system — in the following called *smoothing* — these entropies can be separated into three classes, the elements of each being approximately equal. The three classes are: the von Neumann entropy, the α -entropies with $\alpha < 1$ and the α -entropies with $\alpha > 1$.

To follow their argument, we define ε -smooth Rényi α -Entropies,

$$H_{\alpha}^{\varepsilon}(A)_{\rho} = \begin{cases} \min_{\tilde{\rho}} H_{\alpha}(A)_{\tilde{\rho}} & \text{if } \alpha < 1 \\ \max_{\tilde{\rho}} H_{\alpha}(A)_{\tilde{\rho}} & \text{if } \alpha > 1 \end{cases}, \quad \text{where } 0 \leq \varepsilon < 1$$

and the optimization in each case is over an ε -ball of close states, $\tilde{\rho} \approx_{\varepsilon} \rho$. Note that we smooth in the direction of the von Neumann entropy in both ranges, $\alpha < 1$ and $\alpha > 1$.

These entropies now satisfy the following inequalities. (The proof of these statements can be adapted from results in [RW04, RW05].)

$$\begin{aligned} H_{1/2}^{2\varepsilon+\varepsilon'}(A)_{\rho} - \frac{1}{1-\alpha} \log \frac{2}{\varepsilon^2} &\leq H_{\alpha}^{\varepsilon+\varepsilon'}(A)_{\rho} \leq H_{1/2}^{\varepsilon'}(A)_{\rho} + 2 \log \frac{2}{\varepsilon^2} \quad (\alpha < 1), \\ H_{\infty}^{\varepsilon+\varepsilon'}(A)_{\rho} + \frac{1}{\alpha-1} \log \frac{2}{\varepsilon^2} &\geq H_{\alpha}^{\varepsilon'}(A)_{\rho} \geq H_{\infty}^{\varepsilon'}(A)_{\rho} \quad (\alpha > 1). \end{aligned} \quad (4.2)$$

Note that the deviation terms in α and ε do not depend on properties of the state. Hence, if the entropies are large enough, these terms will be negligible in comparison.

On one hand, the second statement of Eq. (4.2) thus shows that smooth Rényi entropies of order $\alpha > 1$ are well approximated by the smooth Rényi entropy of order ∞ . We choose the *min-entropy*, $H_{\min}(A)_{\rho} = H_{\infty}(A)_{\rho}$, as the representative of his class of Rényi entropies. The choice of the Rényi ∞ -entropy is motivated by its operational interpretation as a guessing probability (see [KRS09] and Section 4.4) as well as the fact that its quantum generalization has a simple form that can be expressed as a *semi-definite program*.

On the other hand, the first statement of Eq. (4.2) implies that the smooth Rényi entropies of order $\alpha < 1$ are well approximated by the smooth Rényi entropy of order $\frac{1}{2}$. We call this entropy the *max-entropy*, $H_{\max}(A)_{\rho} = H_{1/2}(A)_{\rho}$. The choice of this entropy as a representative of the class may seem arbitrary at this point. Indeed, it could be argued that the Rényi entropy of order 0 is also a natural choice, as it characterizes such tasks as the amount of memory needed to store the output of a source perfectly in the *one-shot* setting.¹ However, the choice of $\frac{1}{2}$ is motivated

¹In fact, the initial extension to the fully quantum setting was done for the Rényi entropy of order 0 [Ren05].

by the duality relation of the conditional min- and max-entropies, which holds for this choice of the max-entropy (see [KRS09] and Lemma 4.1).

Conditional Rényi entropies can be defined in various ways. In analogy with the von Neumann entropy, we might be tempted to define $H_\alpha(A|B) = H_\alpha(AB) - H_\alpha(B)$. However, to the best of our knowledge, this definition is not very useful to characterize information theoretic tasks in the *one-shot* setting. In this chapter, we propose a natural generalization of the min-entropy that is motivated by its operational interpretation as a guessing probability. Our generalization of the max-entropy then follows immediately from the duality of the min- and max-entropies.

Quantum generalizations of the min- and max-entropies were first considered by Renner and König [RK05, Ren05] in order to investigate security in quantum cryptography and related tasks, e.g. information reconciliation. They considered a generalization of the Rényi-entropy of order 0 and two variations of the conditional min-entropy as well as a different method of smoothing than the one proposed in this thesis. The smooth entropy framework has been consolidated since and we attempt to summarize the most important results in this and the following chapter.

4.1.1 Main Results

The main result of this chapter is a collection of expressions for the quantum conditional min- and max-entropies.

Result 2 (Expressions for the Min- and Max-Entropy). *Let ρ_{ABC} be a pure quantum state, then*

$$\begin{aligned}
H_{\max}(A|C)_\rho &= \log \min \{ \|Z_C\|_\infty : Z_{AC} \in \mathcal{P}(\mathcal{H}_{AC}) \wedge \rho_{ABC} \leq Z_{AC} \otimes \mathbb{1}_B \} \\
&= \log d_A \max_{\sigma} F^2(\rho_{AC}, \pi_A \otimes \sigma_C) \\
&= \log d_A \max_{B \rightarrow B'B''} \max_{\tau} F^2(\rho_{AB'B''}, \gamma_{AB'} \otimes \tau_{B''}) \\
&= \log \min \{ \text{tr}(\sigma) : \sigma \in \mathcal{P}(\mathcal{H}_B) \wedge \rho_{AB} \leq \mathbb{1}_A \otimes \sigma_B \} \\
&= \min_{\sigma} \log \|\sigma_B^{-1/2} \rho_{AB} \sigma_B^{-1/2}\|_\infty \\
&= \min_{\sigma} \inf \{ \lambda \in \mathbb{R} : \rho_{AB} \leq 2^\lambda \mathbb{1}_A \otimes \sigma_B \} \\
&= -H_{\min}(A|B)_\rho,
\end{aligned}$$

where σ, τ are quantum states and $B \rightarrow B'B''$ is an embedding.

In particular, this result can be interpreted as follows. The expression $\max_{B \rightarrow B'B''} \max_{\tau} F(\rho_{AB'B''}, \gamma_{AB'} \otimes \tau_{B''})$ measures the fidelity with a state that corresponds to an omniscient observer $B = B'B''$ of the system A . Any such observer necessarily controls a system B' that is fully entangled with the system A and may, in addition, control a system B'' that

4. MIN- AND MAX-ENTROPIES

is uncorrelated with A . Moreover, the expression $\max_{\sigma} F(\rho_{AC}, \pi_A \otimes \sigma_C)$ measures the fidelity with a state that corresponds to an ignorant observer of the system A . Such an observer C necessarily holds a state σ_C that is product with the system A . Since these two quantities are equal, we find the following: For any pure state ρ_{ABC} , the marginal ρ_{AB} is close to an omniscient observer B of the system A if and only if ρ_{AC} is close to an ignorant observer C of system A .

4.1.2 Outline

In Section 4.2 we formally introduce the conditional min- and max-entropies and show how they can be expressed as *semi-definite programs*. We also formally introduce the von Neumann entropy. In Section 4.3 we evaluate the conditional min- and max-entropies for classical probability distributions, and in Section 4.4, we explore the interpretation of the min-entropy as a guessing probability. Section 4.5 then discusses various properties of the min- and max-entropy. Most importantly, we give first bounds on the min- and max-entropies and show that these entropies are continuous functions of the state.

4.2 Min- and Max-Entropies

4.2.1 The Min-Entropy

The min-entropy of A conditioned on B has already been introduced in previous chapters as a measure of closeness to a state describing a perfect observer of the A system. We start with alternative definitions of the min- and max-entropy and then develop a variety of alternative expressions.

Definition 4.1. Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. The min-entropy of A conditioned on B of the state ρ_{AB} is

$$H_{\min}(A|B)_{\rho} := \max_{\sigma} \sup \{ \lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_B \}, \quad (4.3)$$

where the maximum is taken over all states $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$.

Note that there exist feasible λ only if $\text{supp}\{\sigma_B\} \supseteq \text{supp}\{\rho_B\}$. However, if this condition on the support is satisfied, there exists a feasible $\lambda_* = -\log \|\sigma_B^{-1/2} \rho_{AB} \sigma_B^{-1/2}\|_{\infty}$ which achieves the supremum. The min-entropy can thus alternatively be written as

$$H_{\min}(A|B)_{\rho} = \max_{\sigma} -\log \|\sigma_B^{-1/2} \rho_{AB} \sigma_B^{-1/2}\|_{\infty}, \quad (4.4)$$

where we use the generalized inverse and the maximum is taken over all $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ with $\text{supp}\{\sigma_B\} \supseteq \text{supp}\{\rho_B\}$.

4.2 Min- and Max-Entropies

We can also reformulate (4.3) as a *semi-definite program* (SDP). (Semi-Definite Programs are introduced in Section 2.3.1.) For this purpose, we include $2^{-\lambda}$ in σ_B and allow the new σ_B to be an arbitrary positive semi-definite operator. The min-entropy is then given by

$$H_{\min}(A|B)_\rho = -\log \min \{ \text{tr}(\sigma) : \sigma \in \mathcal{P}(\mathcal{H}_B) \wedge \rho_{AB} \leq \mathbb{1}_A \otimes \sigma_B \}$$

and the optimization problem thus has an efficient numerical solver. In particular, we consider the SDP for the expression $2^{-H_{\min}(A|B)_\rho}$.

primal problem	dual problem
minimize : $\langle \mathbb{1}_B, \sigma_B \rangle$ subject to : $\mathbb{1}_A \otimes \sigma_B \geq \rho_{AB}$ $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$	maximize : $\langle \rho_{AB}, X_{AB} \rangle$ subject to : $\text{tr}_A(X_{AB}) \leq \mathbb{1}_B$ $X_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$

(4.5)

Clearly, the dual problem has a finite solution; in fact, it is easy to verify that $\langle \rho_{AB}, X_{AB} \rangle \leq \text{tr}(X_{AB}) \leq d_B$. Furthermore, there exists a $\sigma_B > 0$ with $\mathbb{1}_A \otimes \sigma_B > \rho_{AB}$. Hence, strong duality (Theorem 2.6) applies and the primal and dual solution are equivalent.

Let us now investigate the dual problem more closely. We can replace the inequality in the condition $X_B \leq \mathbb{1}_B$ by an equality since adding a positive part to X_{AB} only increases $\langle \rho_{AB}, X_{AB} \rangle$. Hence, X_{AB} can be interpreted as a *Choi-Jamiolkowski* state of a completely positive *unital* map (cf. Section 2.1.5) from $\mathcal{H}_{B'} \cong \mathcal{H}_A$ to \mathcal{H}_B . Let \mathcal{E}^\dagger be that map, then

$$2^{-H_{\min}(A|B)_\rho} = \max_{\mathcal{E}^\dagger} \langle \rho_{AB}, \mathcal{E}^\dagger[\Gamma_{AB'}] \rangle = d_A \max_{\mathcal{E}} \langle \mathcal{E}[\rho_{AB}], \gamma_{AB'} \rangle,$$

where the second maximization is over all *TP-CPMs* \mathcal{E} from B to B' , i.e. all super-operators whose adjoint is completely positive and unital from B' to B . The fully entangled state $\gamma = \Gamma/d_A$ is pure and normalized, hence, we can write [KRS09]

$$H_{\min}(A|B)_\rho = -\log d_A \max_{\mathcal{E}} F^2(\mathcal{E}[\rho_{AB}], \gamma_{AB'}), \quad (4.6)$$

where the maximum is taken over all TP-CPMs from B to B' . (Note that γ is defined as the fully entangled in an arbitrary but fixed basis of $\mathcal{H}_A \cong \mathcal{H}_{B'}$. The expression is invariant under the choice of basis, since the fully entangled states can be converted into each other by a unitary appended to \mathcal{E} .) We write F for the generalized fidelity, which corresponds to the fidelity in this case as γ is normalized.

Alternatively, we can interpret X_{AB} as the Choi-Jamiolkowski state of a TP-CPM map from $\mathcal{H}_{A'} \cong \mathcal{H}_B$ to \mathcal{H}_A . This immediately leads to the relation

$$H_{\min}(A|B)_\rho = -\log d_B \max_{\mathcal{E}} \langle \rho_{AB}, \mathcal{E}[\gamma_{A'B}] \rangle,$$

4. MIN- AND MAX-ENTROPIES

where the maximization is over all TP-CPMs from A' to A .

We may now decompose the TP-CPMs of (4.6) into their Stinespring dilation: an isometry $U : \mathcal{H}_B \rightarrow \mathcal{H}_{B'B''}$ followed by a partial trace over $\mathcal{H}_{B'}$. Uhlmann's theorem now implies that there exists an extension of $\gamma_{AB'}$ to B'' such that $F(U\rho_{AB}U^\dagger, \gamma_{AB'B''}) = F(\mathcal{E}[\rho_{AB}], \gamma_{AB'})$. Since such extensions of a pure state are necessarily of the form $\gamma_{AB'B''} = \gamma_{AB'} \otimes \tau_{B''}$, we recover the following expression for the min-entropy

$$H_{\min}(A|B)_\rho = -\log d_A \max_{B \rightarrow B'B''} \max_{\tau} F^2(\rho_{AB'B''}, \gamma_{AB'} \otimes \tau_{B''}), \quad (4.7)$$

where the maximization is over all isometries from B to $B'B''$ and states $\tau \in \mathcal{S}_=(\mathcal{H}_{B''})$.

Using the expression in (4.7), the min-entropy can be interpreted as a measure of distance to a state describing an observer B that is *omniscient* about A . Such an observer must necessarily hold a state γ that is fully entangled with A and may, in addition, hold an arbitrary state τ that is uncorrelated with A . The min-entropy now evaluates the distance (in terms of the fidelity) of ρ to the closest such state.

Finally, we introduce the quantity $\hat{H}_{\min}(A|B)$, which is a trivial lower bound on $H_{\min}(A|B)$ and is sometimes used instead of $H_{\min}(A|B)$ [Ren05].

$$\hat{H}_{\min}(A|B)_\rho := -\log \|\rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2}\|_\infty \leq H_{\min}(A|B)_\rho.$$

The inequality follows by the choice $\sigma_B = \rho_B$ in (4.4). It has been shown that the smooth versions of $\hat{H}_{\min}(A|B)$ and $H_{\min}(A|B)$ are equivalent up to terms in the smoothing parameter [TSSR11].

4.2.2 The Max-Entropy

We use the following definition of the max-entropy.

Definition 4.2. Let $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$. The max-entropy of A conditioned on B of the state ρ_{AB} is

$$H_{\max}(A|B)_\rho := \max_{\sigma} \log \|\sqrt{\rho_{AB}} \sqrt{\mathbb{1}_A \otimes \sigma_B}\|_1^2, \quad (4.8)$$

where the maximum is taken over all states $\sigma \in \mathcal{S}_\leq(\mathcal{H}_B)$.

Since the maximum is taken for normalized states $\sigma \in \mathcal{S}_=(\mathcal{H}_B)$, we may rewrite this as

$$H_{\max}(A|B)_\rho = \max_{\sigma} \log d_A F(\rho_{AB}, \pi_A \otimes \sigma_B).$$

Contrasting this to the min-entropy in (4.7), the max-entropy can be seen as a measure of proximity of ρ to a state describing an observer B that is ignorant about A . Such an observer necessarily holds a state that is

4.2 Min- and Max-Entropies

product with the state on A and the max-entropy evaluates the fidelity with the closest such state.

Introducing an arbitrary *purification* ρ_{ABC} of ρ_{AB} and applying *Uhlmann's theorem*, we rewrite this as the following optimization problem.

$$2^{H_{\max}(A|B)_\rho} = d_A \max_{\tau} \langle \rho_{ABC} | \tau_{ABC} | \rho_{ABC} \rangle,$$

where τ has the marginal $\tau_{AB} = \pi_A \otimes \sigma_B$ for some $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$. This is the dual problem of the following SDP:

primal problem	dual problem
minimize : μ subject to : $\mu \mathbb{1}_B \geq \text{tr}_A(Z_{AB})$ $Z_{AB} \otimes \mathbb{1}_C \geq \rho_{ABC}$ $Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ $\mu \geq 0$	maximize : $\langle \rho_{ABC}, Y_{ABC} \rangle$ subject to : $\text{tr}_C(Y_{ABC}) \leq \mathbb{1}_A \otimes \sigma_B$ $\text{tr}(\sigma_B) \leq 1$ $Y_{ABC} \in \mathcal{P}(\mathcal{H}_{ABC})$ $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$

Again, it is easy to verify that the dual problem has a finite solution. To see this, note that $\text{tr}(Y) \leq d_A$ due to the constraints in the dual problem, hence, the maximum cannot exceed d_A for normalized states. Moreover, we can easily construct a primal feasible solution with $Z_{AB} \otimes \mathbb{1}_C > \rho_{ABC}$ and $\mu \mathbb{1}_B > Z_B$. Hence, strong duality (Theorem 2.6) applies and the primal and dual solution are equivalent.

The primal problem can be rewritten by noting that the optimization over μ corresponds to evaluating the ∞ -norm of Z_B .

$$H_{\max}(A|B)_\rho = \log \min \left\{ \|Z_B\|_{\infty} : Z_{AB} \otimes \mathbb{1}_C \geq \rho_{ABC}, Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB}) \right\}. \quad (4.9)$$

This can be used to prove upper bounds on the max-entropy. For example, the quantity $\hat{H}_{\max}(A|B)$ — which is sometimes used instead of the max-entropy [Ren05] — is an upper to $H_{\max}(A|B)$.

$$\hat{H}_{\max}(A|B)_\rho := \log \max_{\sigma} \text{tr}(\Pi^{\rho_{AB}} \mathbb{1}_A \otimes \sigma_B) \geq H_{\max}(A|B)_\rho.$$

This follows from (4.9) by the choice $Z_{AB} = \Pi^{\rho_{AB}}$, which is the projector onto the support of ρ_{AB} . Note also that smooth versions of $\hat{H}_{\max}(A|B)$ and $H_{\max}(A|B)$ are equivalent up to terms in the smoothing parameter [TSSR11].

Furthermore, König et al. [KRS09] showed that the max-entropy can be expressed as a min-entropy of the purified state.

Lemma 4.1. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ be pure. Then,*

$$H_{\max}(A|B)_\rho = -H_{\min}(A|C)_\rho.$$

4. MIN- AND MAX-ENTROPIES

Proof. We show that $2^{H_{\max}(A|B)}/d_A = 2^{-H_{\min}(A|C)}/d_A$ using the expression (4.6) for the min-entropy and Def. 4.2 for the max-entropy, i.e. we will show that

$$\max_{\sigma} \|\sqrt{\rho_{AB}}\sqrt{\pi_A \otimes \sigma_B}\|_1 = \max_{C \rightarrow C'C''} \max_{\tau} \|\sqrt{\rho_{AC'C''}}\sqrt{\gamma_{AC'} \otimes \tau_{C''}}\|_1,$$

where the maximization is over quantum states $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ and $\tau_{C''} \in \mathcal{S}_{\leq}(\mathcal{H}_{C''})$ as well as embeddings $C \rightarrow C'C''$.

Due to Uhlmann's theorem (3.3) and the fact that the fidelity cannot decrease under partial trace, we have

$$\begin{aligned} \|\sqrt{\rho_{AB}}\sqrt{\pi_A \otimes \sigma_B}\|_1 &= \max_{C \rightarrow C'C''} \|\sqrt{\rho_{ABC'C''}}\sqrt{\gamma_{AC'} \otimes \varphi_{BC''}}\|_1 \\ &\leq \max_{C \rightarrow C'C''} \max_{\tau} \|\sqrt{\rho_{AC'C''}}\sqrt{\gamma_{AC'} \otimes \tau_{C''}}\|_1, \end{aligned}$$

where $|\varphi\rangle$ is any purification of σ and the fully entangled state $|\gamma\rangle$ purifies the fully mixed state π by definition. Since this holds for every σ , it particularly holds when σ maximizes the fidelity in the definition of the max-entropy. On the other hand,

$$\begin{aligned} \|\sqrt{\rho_{AC'C''}}\sqrt{\gamma_{AC'} \otimes \tau_{C''}}\|_1 &= \max_{|\varphi\rangle} \|\sqrt{\rho_{ABC'C''}}\sqrt{\gamma_{AC'} \otimes \varphi_{BC''}}\|_1 \\ &\leq \max_{\sigma} \|\sqrt{\rho_{AB}}\sqrt{\pi_A \otimes \sigma_B}\|_1. \end{aligned}$$

Since this holds for all embeddings $C \rightarrow C'C''$ and all states τ , it particularly holds for the tuple that achieves the maximum fidelity in (4.6). Thus, we established the equality of the two expressions. \square

4.2.3 The von Neumann Entropy

For completeness, we also define the *von Neumann entropy* for sub-normalized states.

Definition 4.3. Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Then, the von Neumann entropy is

$$H(A|B)_{\rho} := H(\rho_{AB}) - H(\rho_B), \quad \text{where} \quad H(\rho) := -\text{tr}(\rho \log \rho).$$

We may rewrite this as an optimization problem. For any $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ with $\text{tr}(\sigma) \leq \text{tr}(\rho)$, we find

$$\begin{aligned} H(A|B)_{\rho} &= \text{tr}(\rho_{AB}(\mathbb{1}_A \otimes \log \sigma_B - \log \rho_{AB})) - \text{tr}(\rho_B(\log \sigma_B - \log \rho_B)) \\ &\geq \text{tr}(\rho_{AB}(\mathbb{1}_A \otimes \log \sigma_B - \log \rho_{AB})). \end{aligned}$$

We used that the second term, the relative entropy $\text{tr}(\rho_B(\log \sigma_B - \log \rho_B))$, is non-negative thanks to Klein's inequality [Kle31]. Furthermore, the term

4.3 Relation to Classical Entropies

vanishes for the choice $\sigma_B = \rho_B$. If we optimize the expression over all $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ with $\text{tr}(\sigma) \leq \text{tr}(\rho)$, we therefore get

$$H(A|B)_\rho = \max_{\sigma} \text{tr}(\rho_{AB}(\mathbb{1}_A \otimes \log \sigma_B - \log \rho_{AB})). \quad (4.10)$$

We have seen that the min- and max-entropy satisfy a *duality* relation in Lemma 4.1. Here, we show that the von Neumann entropy also satisfies a duality relation. Note that $H(\rho) = \sum_i \lambda_i \log \frac{1}{\lambda_i}$ is a function of the *eigenvalues* $\{\lambda_i\}$ of ρ only. Due to the *Schmidt decomposition* of any pure state ρ_{ABC} , it holds that

$$H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B) = H(\rho_C) - H(\rho_{AC}) = -H(A|C). \quad (4.11)$$

4.3 Relation to Classical Entropies

Here, we evaluate the min- and max-entropy for states on two classical registers, X and Y . Their content is described by a joint probability distribution P_{XY} encoded in the state

$$\rho_{XY} = \sum_{x,y} P_{XY}(x,y) |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y.$$

4.3.1 Classical Min-Entropy

In order to evaluate the min-entropy $H_{\min}(X|Y)_P = H_{\min}(X|Y)_\rho$ for this state, we consider the SDP for the min-entropy. We use the operators

$$\sigma_Y = \sum_y P_{XY}(x_*^y, y) |y\rangle\langle y| \quad \text{and} \quad X_{AB} = \sum_y |x_*^y\rangle\langle x_*^y| \otimes |y\rangle\langle y|,$$

where x_*^y is the x that maximizes $P_{XY}(x, y)$ for a given y . It is easy to verify that σ_Y is primal feasible and X_{AB} is dual feasible. Since they both give the same (upper and lower, respectively) bound on the min-entropy, we get the equality

$$H_{\min}(X|Y)_P = -\log \sum_y P_{XY}(x_*^y, y) = -\log \sum_y P_Y(y) 2^{-H_{\min}(X)_{Py}}, \quad (4.12)$$

where $H_{\min}(X)_{Py} = \log \max_x P_X^y(x)$ is the min-entropy of X evaluated for the conditional probability distribution $P_X^y(x) = P_{XY}(x, y)/P_Y(y)$.

4.3.2 Classical Max-Entropy

For the max-entropy $H_{\max}(X|Y)_P = H_{\max}(X|Y)_\rho$ the calculation is a bit more involved. First, note that we can assume that σ_Y in the optimization

4. MIN- AND MAX-ENTROPIES

of Def. 4.2 is classical in the same basis as ρ_Y . If the fidelity $F(\rho_{XY}, \mathbb{1}_X \otimes \sigma_Y)$ is maximal for a given σ_Y , we can always measure both ρ_Y and σ_Y in the classical basis of ρ_Y . This operation cannot decrease the fidelity due to (3.2) and the measured state σ_Y^* thus achieves the optimum and is of the form

$$\sigma_Y^* = \sum_y Q_Y(y) |y\rangle\langle y|, \quad \text{where } Q_Y \text{ is a probability distribution.}$$

Hence, we maximize over all probability distributions Q_Y so that

$$H_{\max}(X|Y)_P = \max_{Q_Y} \log \left(\sum_{x,y} \sqrt{P_{XY}(x,y)} \sqrt{Q_Y(y)} \right)^2.$$

It is straightforward to verify—for example using the method of Lagrange multipliers—that the optimal Q_Y is proportional to the map $y \mapsto (\sum_x \sqrt{P_{XY}(x,y)})^2$. Hence, if we insert the (normalized) optimal Q_Y into the above equation, we get

$$H_{\max}(X|Y)_P = \log \left(\sum_y \left(\sum_x \sqrt{P_{XY}(x,y)} \right)^2 \right) = \log \sum_y P_Y(y) 2^{H_{\max}(X)_{Py}},$$

where $H_{\max}(X)_{Py} = 2 \log \sum_x \sqrt{P_X^y(x)} = H_{1/2}(X)_{Py}$ is the Rényi entropy of order $1/2$.

4.4 Guessing Probability

The classical min-entropy $H_{\min}(X|Y)_P$ in (4.12) can be interpreted as a *guessing probability*. Consider an observer with access to Y . What is the probability that this observer guesses X correctly, using his optimal strategy? The optimal strategy of the observer is clearly to guess the x with the highest probability conditioned on his observation y . As before, we denote the probability distribution of x conditioned on a fixed y by P_X^y . Then, the (average) guessing probability is given by

$$\sum_y P_Y(y) \max_x P_X^y(x) = 2^{-H_{\min}(X|Y)_P} \quad (4.13)$$

It was noted by König *et. al.* [KRS09] that this interpretation of the min-entropy extends to the case where Y is replaced by a quantum system B and the allowed strategies include any measurement of B .

Consider a *CQ-state* $\rho_{XB} = \sum_x P_X(x) |x\rangle\langle x| \otimes \omega_B^x$ with $\omega^x \in \mathcal{S}_=(\mathcal{H}_B)$. For states of this form, the expression in (4.6) simplifies to

$$\begin{aligned} 2^{-H_{\min}(X|B)_\rho} &= \max_{\mathcal{E}} \langle \Gamma_{XB'} | \sum_x P_X(x) |x\rangle\langle x| \otimes \mathcal{E}[\omega_B^x] | \Gamma_{XB'} \rangle \\ &= \max_{\mathcal{E}} \sum_x P_X(x) \langle x | \mathcal{E}[\omega_B^x] | x \rangle. \end{aligned}$$

4.5 Properties of the Min- and Max-Entropy

This expression reaches its minimum when \mathcal{E} is a measurement, i.e. a map from B to a register $\mathcal{H}_{B'} \cong \mathcal{H}_X$. Moreover, the expression optimizes the probability that $\mathcal{E}[\omega_B^x]$ is mapped to $|x\rangle$ over all such measurement. We can interpret $|x\rangle_{B'}$ as the observer's guess of the value x and thus $2^{-H_{\min}(X|B)_\rho}$ is the guessing probability in the sense described above.¹

4.5 Properties of the Min- and Max-Entropy

Many properties of the min- and max-entropies (and their smooth variants, introduced in Section 5.2) can be derived from properties of the SDP for the min-entropy, Eq. (4.5). Let us consider the functional $\Phi_{A|B} : \rho_{AB} \mapsto 2^{-H_{\min}(A|B)_\rho}$, given by the SDP, which we extend to arbitrary Hermitian arguments $K \in \mathcal{H}(\mathcal{H}_{AB})$. We use the shorthand notation $\Phi \equiv \Phi_{A|B}$ in the following. We restate the SDP here for the convenience of the reader.

primal problem	dual problem
minimize : $\langle \mathbb{1}_B, \sigma_B \rangle$ subject to : $\mathbb{1}_A \otimes \sigma_B \geq K_{AB}$ $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$	maximize : $\langle K_{AB}, X_{AB} \rangle$ subject to : $\text{tr}_A(X_{AB}) \leq \mathbb{1}_B$ $X_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$

We may express the min- and max-entropies of A conditioned on B of a pure state $\rho \in \mathcal{S}_\leq(\mathcal{H}_{ABC})$ as follows. The min-entropy is given as $H_{\min}(A|B)_\rho = -\log \Phi_{A|B}(\rho_{AB})$ and the max-entropy as $H_{\max}(A|B)_\rho = \log \Phi_{A|C}(\rho_{AC})$, by duality. The most important properties of the functional Φ are listed in Table 4.1.

Some properties can be readily verified by close inspection of the SDP and we will only provide their proof when necessary.

Proof of Property iii. The inequality follows from the fact that dual feasibility of X_{AB} is independent of the argument of Φ . Hence, if X_{AB} is optimal for $K + L$, we have $\Phi(K + L) = \langle K, X \rangle + \langle L, X \rangle \leq \Phi(K) + \Phi(L)$.

For equality, consider the optimal primal for $K + L$, which we denote σ_B . Then, $\Pi^{L_B} \sigma_B \Pi^{L_B}$ is primal feasible for L and $\Pi^{K_B} \sigma_B \Pi^{K_B}$ is primal feasible for K . Hence, if these projectors are orthogonal, we have

$$\Phi(K) + \Phi(L) = \text{tr}((\Pi^{K_B} + \Pi^{L_B})\sigma_B) \leq \text{tr}(\sigma) = \Phi(K + L). \quad \square$$

Proof of Property vi. Let σ_B be the optimal primal for M . Then, $\mathcal{E}[\sigma]$ is primal feasible for $\mathcal{E}[M]$ since

$$\mathbb{1}_A \otimes \sigma_B \geq M_{AB} \implies \mathbb{1}_A \otimes \mathcal{E}[\sigma_B] \geq \mathcal{E}[M_{AB}].$$

¹Note that the special case when B is classical is recovered by setting $\mathcal{E} : |y\rangle \mapsto |x_*^y\rangle$.

4. MIN- AND MAX-ENTROPIES

Let $K, L \in \mathcal{H}(\mathcal{H}_{AB})$ and $M \in \mathcal{P}(\mathcal{H}_{AB})$.

- i. Multiplication with scalar: Let $\lambda \geq 0$. Then, $\Phi(\lambda K) = \lambda \Phi(K)$.
 - ii. Monotonicity: $K \geq L \implies \Phi(K) \geq \Phi(L)$.
 - iii. Sub-Additivity: $\Phi(K + L) \leq \Phi(K) + \Phi(L)$. Furthermore, equality holds if K_B and L_B have orthogonal support.
 - iv. Convexity: $\Phi(\lambda K + (1 - \lambda)L) \leq \lambda \Phi(K) + (1 - \lambda)\Phi(L)$ for $\lambda \in [0, 1]$.
 - v. Invariance under Isometries: Φ is invariant under local isometries on the A or B system. Namely, $\Phi_{A'|B'}((U \otimes V)K(U^\dagger \otimes V^\dagger)) = \Phi(K)$, where $U : \mathcal{H}_A \rightarrow \mathcal{H}_{A'}$ and $V : \mathcal{H}_B \rightarrow \mathcal{H}_{B'}$ are isometries.
 - vi. Data Processing: Let \mathcal{E} be trace non-increasing CPM from B to B' and let \mathcal{F} be a sub-unital CPM from A to A' ($\mathcal{F}[\mathbb{1}_A] \leq \mathbb{1}_{A'}$). Then, we have $\Phi_{A|B'}(\mathcal{E}[M]) \leq \Phi(M)$ and $\Phi_{A'|B}(\mathcal{F}[M]) \leq \Phi(M)$.
 - vii. Data Processing on Extension: Let \mathcal{E} be a trace non-increasing CPM from C to C' . Then, $\Phi(\mathcal{E}[M]) \leq \Phi(M)$.
 - viii. Tensor Product: $\Phi_{A_1 A_2 | B_1 B_2}(L \otimes K) = \Phi_{A_1 | B_1}(L) \Phi_{A_2 | B_2}(K)$.
-

Table 4.1: *Properties of Φ .*

Moreover, $\text{tr}(\mathcal{E}[\sigma]) \leq \text{tr}(\sigma)$, concluding the proof of the first statement. To prove the second statement, note that σ is primal feasible for $\mathcal{F}[M]$ since

$$\mathbb{1}_A \otimes \sigma_B \geq M_{AB} \implies \mathbb{1}_{A'} \otimes \sigma_B \geq \mathcal{F}[\mathbb{1}_A] \otimes \sigma_B \geq \mathcal{F}[M_{AB}],$$

where we applied the defining property of sub-unital maps. \square

Proof of Property vii. Lemma A.1 establishes that $\text{tr}_C(\mathcal{E}[M]) \leq \text{tr}_C(M)$. The property then follows from monotonicity. \square

Proof of Property viii. If σ_{B_1} is primal optimal for L and σ_{B_2} is primal optimal for K , then $\sigma_{B_1} \otimes \sigma_{B_2}$ is primal feasible for $L \otimes K$. Moreover, if $X_{A_1 B_1}$ is dual optimal for L and $X_{A_2 B_2}$ is dual optimal for K , then $X_{A_1 B_1} \otimes X_{A_2 B_2}$ is dual feasible for $L \otimes K$. Hence, the equality follows from properties of the inner product. \square

4.5.1 First Bounds and Order

We first show the following technical lemma.

Lemma 4.2. *Let $K \in \mathcal{H}(\mathcal{H}_{AB})$, $d_A^* = \text{rank}\{\text{tr}_B\{K\}_+\} \leq d_A$ and $d_B^* = \text{rank}\{\text{tr}_A\{K\}_+\} \leq d_B$. Then,*

$$\frac{1}{d_A^*} \text{tr}\{K\}_+ \leq \Phi(K) \leq \min\{d_A^*, d_B^*\} \text{tr}\{K\}_+.$$

4.5 Properties of the Min- and Max-Entropy

Proof. Using the eigenvalue decomposition $K = \sum_i \lambda_i |\varphi^i\rangle\langle\varphi^i|$ and Properties i. to iii., we find

$$\Phi(K) \leq \Phi\left(\sum_{i:\lambda_i>0} \lambda_i \varphi^i\right) \leq \sum_{i:\lambda_i>0} \lambda_i \Phi(\varphi^i).$$

Since the φ^i are normalized and pure, we have, $\varphi^i \leq \mathbb{1}_A \otimes \Pi^{\varphi_B^i}$ and, thus, $\sigma_B = \Pi^{\varphi_B^i}$ is primal feasible. This means that $\Phi(\varphi^i) \leq \text{tr}(\Pi^{\varphi_B^i}) = \text{tr}(\Pi^{\varphi_A^i}) \leq \min\{d_A^*, d_B^*\}$, where the equality follows from the Schmidt decomposition and the latter inequality from

$$\text{supp}\{\varphi_B^i\} \subseteq \text{supp}\{\text{tr}_A\{K\}_+\} \quad \text{and} \quad \text{supp}\{\varphi_A^i\} \subseteq \text{supp}\{\text{tr}_B\{K\}_+\}.$$

This concludes the proof of the upper bound.

On the other hand, we use that

$$X_{AB} = \frac{1}{d_A^*} \Pi^{\text{tr}_B\{K\}_+} \otimes \Pi^{\text{tr}_A\{K\}_+} \geq \frac{1}{d_A^*} \Pi^{\{K\}_+}.$$

is dual feasible. Hence, $\Phi(K) \geq \text{tr}(K_{AB} X_{AB}) \geq \frac{1}{d_A^*} \text{tr}\{K\}_+.$ \square

Using these properties, we now establish bounds on the conditional min- and max-entropies in terms of the support of the marginals of the state. This also establishes that — for normalized states — the min-entropy is always smaller than the von Neumann entropy, which is in turn smaller than the max-entropy.

Proposition 4.3. *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, $d_A^* = \text{rank}\{\rho_A\}$, $d_B^* = \text{rank}\{\rho_B\}$ and $d_{AB}^* = \text{rank}\{\rho_{AB}\}$. Then, using $t = 1/\text{tr}(\rho)$, the following bounds hold:*

$$\begin{aligned} -\log \min\{d_A^*, d_B^*\} &\leq H_{\min}(A|B)_\rho - \log t \\ &\leq t H(A|B)_\rho \\ &\leq H_{\max}(A|B)_\rho + \log t \leq \log \min\{d_A^*, d_{AB}^*\}. \end{aligned}$$

Proof. The lower bound on the min-entropy follows directly from Lemma 4.2. To get the second inequality, we consider the tuple $\{\sigma_B, \lambda\}$ that optimizes the min-entropy in Def. 4.1. Hence, $\mathbb{1}_A \otimes \text{tr}(\rho)\sigma_B \geq \text{tr}(\rho)2^\lambda \rho_{AB}$. We now plug $\text{tr}(\rho)\sigma_B$ into (4.10), providing us with a lower bound on the von Neumann entropy.

$$\begin{aligned} H(A|B)_\rho &\geq \text{tr}(\rho_{AB}(\mathbb{1}_A \otimes \log(\text{tr}(\rho)\sigma_B) - \log \rho_{AB})) \\ &\geq \text{tr}(\rho_{AB}(\log(\text{tr}(\rho)2^\lambda \rho_{AB}) - \log \rho_{AB})) \\ &= \text{tr}(\rho)(\lambda + \log \text{tr}(\rho)), \end{aligned}$$

4. MIN- AND MAX-ENTROPIES

where we used the operator monotonicity of the log function (cf. Section 2.3.2) to get the second inequality. This establishes the second inequality of the statement of the lemma.

The third and forth inequality follow symmetrically from the first two inequalities applied to the marginal state ρ_{AC} of any purification $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ of ρ together with the duality relations $H_{\max}(A|B) = -H_{\min}(A|C)$ and $H(A|B) = -H(A|C)$. Note also that $d_C^* = d_{AB}^*$ due to the Schmidt decomposition. \square

For example, the above lemma implies that the min- and max-entropy of normalized pure states is at most zero and that the min-entropy cannot exceed the max-entropy for normalized states.

4.5.2 Continuity

For classical-quantum states, the operational interpretation of the conditional min-entropy as a guessing probability (cf. [KRS09]) already implies its continuity in the state. To see this, note that a discontinuity in the guessing probability could be detected experimentally using a fixed number of trials (the number depending only on the required precision), hence giving us the means to distinguish between arbitrarily close states for a cost (in terms of the number of trials) independent of their distance. For sufficiently close states, this would contradict the upper bound on the distinguishing advantage, Eq. (3.1). Here, we make this statement more precise.

Proposition 4.4. *Let $\rho, \tau \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\delta := D(\rho, \tau)$, then*

$$|H_{\min}(A|B)_{\rho} - H_{\min}(A|B)_{\tau}| \leq \delta \frac{d_A \min\{d_A, d_B\}}{\ln 2 \cdot \min\{\text{tr}(\rho), \text{tr}(\sigma)\}}.$$

Proof. We use continuity of the functional Φ to obtain

$$\begin{aligned} \Phi(\tau) &= \Phi(\rho + (\tau - \rho)) \leq \Phi(\rho) + \Phi(\tau - \rho) \\ &\leq \Phi(\rho) + \min\{d_A, d_B\} \text{tr}\{\tau - \rho\}_+ \leq \Phi(\rho) + \min\{d_A, d_B\} \delta. \end{aligned} \quad (4.14)$$

Note that $\Phi > 0$ for all states in $\mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Taking the logarithm and using the bound $\ln(a + x) \leq \ln a + \frac{x}{a}$, we find

$$\log \Phi(\tau) - \log \Phi(\rho) \leq \delta \frac{\min\{d_A, d_B\}}{\ln 2 \cdot \Phi(\rho)} \leq \delta \frac{d_A \min\{d_A, d_B\}}{\ln 2 \cdot \text{tr}(\rho)}.$$

The same argument also applies on exchange of ρ and τ and we obtain the statement of the lemma by substituting $H_{\min}(A|B)_{\rho} = -\log \Phi(\rho)$. \square

4.5 Properties of the Min- and Max-Entropy

Remark 4.5. *The above result is tight in the following sense: Consider a system with Hilbert spaces \mathcal{H}_A and $\mathcal{H}_B = \mathcal{H}_{A'} \oplus \mathcal{H}_{B'}$, where $\mathcal{H}_{A'} \cong \mathcal{H}_A$. Let γ_{AB} be the normalized fully entangled state on $\mathcal{H}_A \otimes \mathcal{H}_{A'}$ and $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_{B'})$ be orthogonal to γ_B . The choice $\rho = \pi_A \otimes \sigma_B$ and $\tau = \rho + \delta\gamma$ for some small $\delta > 0$ leads to $D(\rho, \tau) = \delta$,*

$$\Phi(\rho) = \frac{\text{tr}(\rho)}{d_A} \quad \text{and} \quad \Phi(\tau) = \Phi(\rho) + \delta \min\{d_A, d_B\}.$$

Taking the logarithm (for small δ) leads to

$$\log \Phi(\tau) - \log \Phi(\rho) \approx \delta \frac{\min\{d_A, d_B\}}{\ln 2 \cdot \Phi(\rho)} = \delta \frac{d_A \min\{d_A, d_B\}}{\ln 2 \cdot \text{tr}(\rho)}.$$

Proposition 4.4 implies that the conditional min-entropy is uniformly (Lipschitz) continuous on the set of normalized states and in any ε -ball. Since $D(\rho, \tau) \leq P(\rho, \tau)$, the bound also holds for $\delta = P(\rho, \tau)$. Due to the duality between min- and max-entropy, Lipschitz continuity in $P(\rho, \tau)$ also follows for the max-entropy.

4.5.3 Conditioning on Classical Information

Let us, more generally, consider a general state between two quantum systems, A and B , and a classical register, K . Such a state has the form

$$\rho_{ABK} = \sum_k p_k |k\rangle\langle k| \otimes \tau_{AB}^k, \quad \text{where } \tau^k \in \mathcal{S}_{\leq}(\mathcal{H}_{AB}) \quad (4.15)$$

and $\{p_k\}$ is a probability distribution.

Proposition 4.6. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABK})$ be of the form (4.15). Then,*

$$\begin{aligned} H_{\min}(A|BK)_{\rho} &= -\log \left(\sum_k p_k 2^{-H_{\min}(A|B)_{\tau^k}} \right) \quad \text{and} \\ H_{\max}(A|BK)_{\rho} &= \log \left(\sum_k p_k 2^{H_{\max}(A|B)_{\tau^k}} \right). \end{aligned}$$

Proof. The first statement directly from the sub-additivity of Φ , Property ii., and the fact that the marginal states on the conditioning system, $\tau_B^k \otimes |k\rangle\langle k|$, are orthogonal. Thus,

$$\Phi_{A|BK}(\rho) = \sum_k p_k \Phi_{A|BK}(\tau^k \otimes |k\rangle\langle k|) = \sum_k p_k \Phi_{A|B}(\tau^k).$$

Note that we can remove the trivial register K in the last expression since Φ is invariant under the isometry $\tau \mapsto \tau \otimes |k\rangle\langle k|$.

4. MIN- AND MAX-ENTROPIES

The second statement follows by the duality relation of the min- and max-entropy (Lemma 4.1). We consider purifications of τ^k on C and introduce a purification of ρ as

$$|\rho\rangle_{ABCKK'} = \sum_k \sqrt{p_k} |\tau^k\rangle_{ABC} \otimes |k\rangle_K \otimes |k\rangle_{K'}, \quad (4.16)$$

where $\mathcal{H}_{K'} \cong \mathcal{H}_K$. Using this state and the fact that its marginal $\rho_{ACK'}$ is of the form (4.15), we find

$$\begin{aligned} H_{\max}(A|BK)_\rho &= -H_{\min}(A|CK')_\rho = \log \left(\sum_k p_k 2^{-H_{\min}(A|C)_{\tau^k}} \right) \\ &= \log \left(\sum_k p_k 2^{H_{\max}(A|B)_{\tau^k}} \right). \end{aligned} \quad \square$$

4.5.4 Concavity of the Max-Entropy

The following bounds follow directly from the above lemma and the *concavity* of the log function. For a state of the form (4.15), we have

$$\begin{aligned} H_{\min}(A|BK)_\rho &\leq \sum_k p_k H_{\min}(A|B)_{\tau^k} \quad \text{and} \\ H_{\max}(A|BK)_\rho &\geq \sum_k p_k H_{\max}(A|B)_{\tau^k}. \end{aligned}$$

Moreover, $H_{\max}(A|B)_\rho = \log \Phi_{A|CKK'}(\rho) \geq \log \Phi_{A|CK'}(\rho) = H_{\max}(A|BK)_\rho$ due to *data processing* of Φ applied to the purification of ρ given in (4.16). Together with the fact that $\rho_{AB} = \sum_k p_k \tau_{AB}^k$, this implies that the max-entropy is a concave function of the state.

Furthermore, we have seen that $\Phi_{A|B}$ (which corresponds to the guessing probability if A is classical) is convex in the state. However, numerical evidence suggests that the min-entropy itself is neither *concave* nor *convex* in the state.

Chapter 5

Smooth Entropies

Smooth entropies are defined as optimizations of the min- and max-entropy over a set of close states. In this chapter, based on [TCR10], we propose that this closeness should be measured in terms of the purified distance. This endows the smooth entropies with many useful properties, for example invariance under isometries and a duality relation. Among many other properties, we also show that the smooth entropies satisfy a *data processing* inequality and chain rules.

5.1 Introduction and Related Work

Generally, smooth entropies are defined as optimizations of the underlying entropy over a ball of states close to the state under consideration. In particular, the *smooth min- and max-entropies* are defined as

$$H_{\min}^{\varepsilon}(A|B)_{\rho} = \max_{\tilde{\rho}} H_{\min}(A|B)_{\tilde{\rho}} \quad \text{and} \quad H_{\max}^{\varepsilon}(A|C)_{\rho} = \min_{\tilde{\rho}} H_{\max}(A|C)_{\tilde{\rho}},$$

where the optimization is over an ε -ball of states $\tilde{\rho}$ close to ρ . The conditional smooth entropies for quantum states are first introduced by Renner [Ren05]. In this work, various properties of the smooth entropies are shown, among them chain rules and various data processing inequalities.

It is important to note that the metric used in [Ren05] to define the ε -ball was a generalization of the trace distance to sub-normalized states. Subsequently, we found that a metric based on the fidelity would be more suitable [Wul08]. Using such a metric, we can always find extensions, due to Uhlmann's theorem, that are as close as their marginals. This property does not hold for the trace distance metric used in [Ren05]. Note also that Renner used a max-entropy based on the Rényi entropy of order 0 as a basis for the smooth max-entropy and was thus unable to harness the duality relation, which simplifies many proofs.

5. SMOOTH ENTROPIES

5.1.1 Main Contributions

Using the *purified distance* as a metric, we are able to show many properties of the smooth entropies in a much more direct way as compared to [Ren05]. In particular, this smoothing allows us extend the *duality* relation of the min- and max-entropy (cf. Lemma 4.1) to ε -smooth entropies.

Result 3 (Duality Relation). *For any $\varepsilon \geq 0$ and ρ_{ABC} pure, we have*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} = -H_{\max}^{\varepsilon}(A|C)_{\rho}.$$

Another result we want to highlight is a *data processing* inequality. It states that the uncertainty about the A system gets at most larger when we condition on less side information.

Result 4 (Data Processing Inequality). *For any $\varepsilon \geq 0$, we have*

$$H_{\min}^{\varepsilon}(A|BC)_{\rho} \leq H_{\min}^{\varepsilon}(A|B)_{\rho} \quad \text{and} \quad H_{\max}^{\varepsilon}(A|BC)_{\rho} \leq H_{\max}^{\varepsilon}(A|B)_{\rho}.$$

This is a special case of a more general data processing inequality that holds for any *TP-CPM* applied to the system we condition on.

This section also includes a variety of other results. For example, various properties of smooth entropies of classical *registers* are discussed and should be very helpful in many applications. For example, we show that the smoothing on registers can be restricted to valid classical states without loss of generality. This implies that the smooth entropies are also well-defined in their classical limit.

5.1.2 Outline

In Section 5.2, we introduce the smooth min- and max-entropies and show several properties of the smoothing. Section 5.3 discusses some properties of the smooth entropies, including their invariance under isometries and a duality relation. One of the most striking properties, a generalized data processing inequality, is shown in Section 5.4. Then, Section 5.5 explores the properties of the smooth entropies of classical registers. Finally, Section 5.6 summarizes some chain rules that were recently proven for the smooth entropies.

5.2 Smooth Min- and Max-Entropies

The smooth entropies of a state ρ are defined as optimization over the min- and max-entropies of states $\tilde{\rho} \approx_{\varepsilon} \rho$, i.e. over states that are close to ρ . Here, we define the smooth min- and max-entropies and explore some properties of the smoothing. In particular, we show that our definition of smoothing allows to extend the duality relation of the min- and max-entropy to the smooth entropies.

5.2.1 The Smooth Entropies

We introduce sets of ε -close states that will then later be used to define the smooth entropies.

Definition 5.1 (ε -Ball). *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $0 \leq \varepsilon < \sqrt{\text{tr}(\rho)}$. We define the ε -ball of operators on \mathcal{H} around ρ as*

$$\mathcal{B}^{\varepsilon}(\mathcal{H}; \rho) := \{\tau \in \mathcal{S}_{\leq}(\mathcal{H}) : P(\tau, \rho) \leq \varepsilon\}.$$

Furthermore, we define the ε -ball of pure states around ρ as $\mathcal{B}_{\text{p}}^{\varepsilon}(\mathcal{H}; \rho) := \{\tau \in \mathcal{B}^{\varepsilon}(\mathcal{H}; \rho) : \text{rank}\{\tau\} = 1\}$.

For the remainder of this chapter, we will assume that ε is sufficiently small so that $\varepsilon < \sqrt{\text{tr}(\rho)}$ is always satisfied. Furthermore, if the Hilbert space used is obvious from context, we will omit it and simply use the notation $\mathcal{B}^{\varepsilon}(\rho)$. We now list some properties of this ε -ball that come in addition to the properties of the underlying purified distance metric.

- i. The set $\mathcal{B}^{\varepsilon}(\mathcal{H}; \rho)$ is compact and convex.

Proof. The set is closed and bounded, hence compact. For convexity, we require that, for any $\lambda \in [0, 1]$ and $\sigma, \tau \in \mathcal{B}^{\varepsilon}(\rho)$, the state $\omega := \lambda\sigma + (1 - \lambda)\tau$ is also in $\mathcal{B}^{\varepsilon}(\rho)$. We define $\hat{\omega} = \omega \oplus (1 - \text{tr} \omega)$ and analogously $\hat{\rho}$, $\hat{\sigma}$ and $\hat{\tau}$. Thus, $\hat{\omega} = \lambda\hat{\sigma} + (1 - \lambda)\hat{\tau}$ by linearity. By assumption, we have $F(\hat{\sigma}, \hat{\rho}) \geq \sqrt{1 - \varepsilon^2}$ and $F(\hat{\tau}, \hat{\rho}) \geq \sqrt{1 - \varepsilon^2}$. We use the concavity of the fidelity (cf. [NC00], Section 9.2.2) to find

$$\begin{aligned} P(\omega, \rho) &= \sqrt{1 - F(\hat{\omega}, \hat{\rho})^2} \\ &= \sqrt{1 - F(\lambda\hat{\sigma} + (1 - \lambda)\hat{\tau}, \hat{\rho})^2} \\ &\leq \sqrt{1 - (\lambda F(\hat{\sigma}, \hat{\rho}) + (1 - \lambda)F(\hat{\tau}, \hat{\rho}))^2} \leq \varepsilon. \end{aligned}$$

Therefore, $\omega \in \mathcal{B}^{\varepsilon}(\rho)$, as required. \square

- ii. Normalized states in $\mathcal{B}^{\varepsilon}(\mathcal{H}; \rho)$ are not distinguishable from a normalized state ρ with probability more than $\frac{1}{2}(1 + \varepsilon)$.

Proof. By Proposition 3.3, $\tau \in \mathcal{B}^{\varepsilon}(\rho)$ implies $D(\tau, \rho) \leq P(\tau, \rho) \leq \varepsilon$. The statement then follows from (3.1). \square

- iii. The ball grows monotonically in the smoothing parameter ε , namely $\varepsilon < \varepsilon' \implies \mathcal{B}^{\varepsilon}(\mathcal{H}; \rho) \subset \mathcal{B}^{\varepsilon'}(\mathcal{H}; \rho)$. Furthermore, $\mathcal{B}^0(\mathcal{H}; \rho) = \{\rho\}$.

The *smooth entropies* are now defined as follows.

5. SMOOTH ENTROPIES

Definition 5.2 (Smooth Entropies). *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\varepsilon \geq 0$. Then, we define the ε -smooth min- and max-entropies of A conditioned on B of the state ρ as*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} := \max_{\tilde{\rho}} H_{\min}^{\varepsilon}(A|B)_{\tilde{\rho}} \quad \text{and} \\ H_{\max}^{\varepsilon}(A|B)_{\rho} := \min_{\tilde{\rho}} H_{\max}^{\varepsilon}(A|B)_{\tilde{\rho}},$$

where the optimization is over all states $\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho_{AB})$ in both cases.

Note that the extrema can be achieved due to compactness of the ε -ball (cf. Property i.). We usually use $\tilde{\rho}_{AB}$ to denote the state that achieves the extremum, e.g., for the min-entropy, there exists a state $\tilde{\rho}_{AB} \in \mathcal{B}^{\varepsilon}(\rho_{AB})$ such that $H_{\min}^{\varepsilon}(A|B)_{\rho} = H_{\min}^{\varepsilon}(A|B)_{\tilde{\rho}}$. The state $\tilde{\rho}$ is ε -indistinguishable from ρ in the sense described in Property ii. Moreover, the smooth min-entropy is monotonically increasing in ε due to Property iii. of the ε -ball and we recover the non-smooth min-entropy in the limit $\varepsilon = 0$, i.e. $H_{\min}^0(A|B)_{\rho} = H_{\min}(A|B)_{\rho}$. Similarly, the smooth max-entropy is monotonically decreasing in ε and $H_{\max}^0(A|B)_{\rho} = H_{\max}(A|B)_{\rho}$.

If ρ_{ABC} is normalized, the optimization problems defining the smooth min- and max-entropies can be formulated as SDPs. To see this, note that the restrictions on the smoothed state $\tilde{\rho}$ are linear in the purification ρ_{ABC} of ρ . In particular, consider the condition $P(\rho, \tilde{\rho}) \leq \varepsilon$ on $\tilde{\rho}$, or, equivalently, $F^2(\rho, \tilde{\rho}) \geq 1 - \varepsilon^2$. If ρ_{ABC} is normalized, then the squared fidelity can be expressed as $F^2(\rho, \tilde{\rho}) = \text{tr}(\rho_{ABC} \tilde{\rho}_{ABC})$.

We give the primal of the SDP for $2^{-H_{\min}^{\varepsilon}(A|B)_{\rho}}$ as an example. This SDP is parametrized by an (arbitrary) purification $\rho \in \mathcal{S}_{=}(\mathcal{H}_{ABC})$.

primal problem

$$\begin{aligned} \text{minimize : } & \langle \mathbb{1}_B, \sigma_B \rangle \\ \text{subject to : } & \mathbb{1}_A \otimes \sigma_B \geq \text{tr}_C(\tilde{\rho}_{ABC}) \\ & \text{tr}(\tilde{\rho}_{ABC}) \leq 1 \\ & \text{tr}(\tilde{\rho}_{ABC} \rho_{ABC}) \geq 1 - \varepsilon^2 \\ & \tilde{\rho}_{ABC} \in \mathcal{P}(\mathcal{H}_{ABC}) \\ & \sigma_B \in \mathcal{P}(\mathcal{H}_B) \end{aligned}$$

This program allows us to efficiently compute the smooth min-entropy.

5.2.2 Remarks on Smoothing

For both the smooth min- and max-entropy, we can restrict the optimization in Def. 5.2 to states in the support of $\rho_A \otimes \rho_B$.

Proposition 5.1. *For any state $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, there exist optimal states $\tilde{\rho}, \bar{\rho} \in \mathcal{B}^{\varepsilon}(\rho_{AB})$ in $\text{supp}\{\rho_A\} \otimes \text{supp}\{\rho_B\}$ such that $H_{\min}^{\varepsilon}(A|B)_{\rho} = H_{\min}^{\varepsilon}(A|B)_{\tilde{\rho}}$ and $H_{\max}^{\varepsilon}(A|B)_{\rho} = H_{\max}^{\varepsilon}(A|B)_{\bar{\rho}}$.*

5.2 Smooth Min- and Max-Entropies

Proof. Let ρ_{ABC} be any purification of ρ on C , where $\mathcal{H}_C \cong \mathcal{H}_{AB}$. Moreover, let $\Pi = \Pi^{\rho_A} \otimes \Pi^{\rho_B}$ be the projector onto the support of $\rho_A \otimes \rho_B$. Recall that we can express the min- and max-entropies in terms of the functional Φ , i.e.

$$H_{\min}(A|B)_\rho = -\log \Phi(\rho) \quad \text{and} \quad H_{\max}(A|B)_\rho = \log \Phi_{A|C}(\rho). \quad (5.1)$$

For the min-entropy, first consider any state $\tilde{\rho}' \in \mathcal{B}^\varepsilon(\mathcal{H}_{AB}, \rho_{AB})$ that achieves the maximum in Def. 5.2, i.e. $H_{\min}^\varepsilon(A|B)_\rho = -\log \Phi(\tilde{\rho}')$. We have $\Phi(\tilde{\rho}') \geq \Phi(\Pi\tilde{\rho}'\Pi)$ due to the data processing property of Φ (Property vi.). This implies that projecting onto Π will not decrease the min-entropy. Moreover, since $\rho = \Pi\rho\Pi$, we find $P(\Pi\tilde{\rho}'\Pi, \rho) \leq P(\tilde{\rho}', \rho) \leq \varepsilon$ due to the monotonicity of the purified distance under trace non-increasing maps. Hence, $\tilde{\rho} = \Pi\tilde{\rho}'\Pi \in \mathcal{B}^\varepsilon(\rho_{AB})$ and, thus, necessarily $H_{\min}(A|B)_{\tilde{\rho}} = H_{\min}^\varepsilon(A|B)_\rho$.

For the max-entropy, consider a state $\tilde{\rho}' \in \mathcal{B}^\varepsilon(\rho_{ABC})$ whose marginal achieves the minimum in Def. 5.2. (Note that every $\tilde{\rho}_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ has a purification in this ε -ball around ρ_{ABC} due to Theorem 3.5.) Thus,

$$H_{\max}^\varepsilon(A|B)_\rho = \log \Phi_{A|C}(\tilde{\rho}') \geq \log \Phi_{A|C}(\Pi^{\rho_A} \tilde{\rho}' \Pi^{\rho_A}) \geq \log \Phi_{A|C}(\Pi\tilde{\rho}'\Pi).$$

Here, the first inequality follows from the data processing property of Φ (property v.) and the second one from data processing on the purifying system (Property vii.). Since $\tilde{\rho} = \Pi\tilde{\rho}'\Pi \in \mathcal{B}^\varepsilon(\rho_{ABC})$, we get $H_{\max}(A|B)_{\tilde{\rho}} = H_{\max}^\varepsilon(A|B)_\rho$. \square

Note that these optimal states are not necessarily normalized. In fact, it is in general not possible to find a normalized state in the support of $\rho_A \otimes \rho_B$ that achieves the optimum. However, if ρ is normalized, we can always find normalized optimal states if we embed the systems A and B into large enough Hilbert spaces that allow smoothing outside the support of $\rho_A \otimes \rho_B$. For the min-entropy, this is intuitively true since adding weight in a space orthogonal to A , if sufficiently diluted, will neither effect the min-entropy nor the purified distance. For the max-entropy, the result follows from the duality of the entropies.

Lemma 5.2. *Let $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$. Then, there exists an embedding from \mathcal{H}_A to $\mathcal{H}_{A'}$ and a normalized state $\hat{\rho}_{A'B} \in \mathcal{B}^\varepsilon(\rho_{A'B})$ such that $H_{\min}(A'|B)_{\hat{\rho}} = H_{\min}^\varepsilon(A|B)_\rho$. Moreover, there exist embeddings \mathcal{H}_A to $\mathcal{H}_{A'}$ and \mathcal{H}_B to $\mathcal{H}_{B'}$ and a state $\hat{\rho}_{A'B'} \in \mathcal{B}^\varepsilon(\rho_{A'B'})$ such that $H_{\max}(A'|B')_{\hat{\rho}} = H_{\max}^\varepsilon(A|B)_\rho$.*

Proof. To prove the property for the min-entropy, let $\{\tilde{\rho}_{AB}, \sigma_B\}$ be such that they maximize the smooth min-entropy $\lambda = H_{\min}^\varepsilon(A|B)_\rho$, i.e. we have $\tilde{\rho}_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_B$. Then we introduce an auxiliary Hilbert space $\mathcal{H}_{\bar{A}}$ with dimension $d_{\bar{A}}$ to be defined and embed \mathcal{H}_A into $\mathcal{H}_{A'} \cong \mathcal{H}_A \oplus \mathcal{H}_{\bar{A}}$. The state

5. SMOOTH ENTROPIES

$\hat{\rho}_{A'B} = \tilde{\rho}_{AB} \oplus (1 - \text{tr } \tilde{\rho}) \pi_{\bar{A}} \otimes \sigma_B$, where $\pi_{\bar{A}} = \mathbb{1}_{\bar{A}}/d_{\bar{A}}$, satisfies the required conditions. We have

$$\hat{\rho}_{A'B} = \tilde{\rho}_{AB} \oplus (1 - \text{tr } \tilde{\rho}) \pi_{\bar{A}} \otimes \sigma_B \leq 2^{-\lambda} (\mathbb{1}_A \oplus \mathbb{1}_{\bar{A}}) \otimes \sigma_B$$

if $2^\lambda(1 - \text{tr } \tilde{\rho}) \leq 2^\lambda \varepsilon \leq d_{\bar{A}}$, i.e. if $d_{\bar{A}}$ is chosen large enough. This implies that the $H_{\min}^\varepsilon(A'|B)_{\hat{\rho}}$ of $\hat{\rho}$ is at least λ . Moreover, $F(\hat{\rho}, \rho) = F(\tilde{\rho}, \rho)$ is not affected by adding weight into an orthogonal subspace.

The equivalent statement for the max-entropy follows by duality. For any purification ρ_{ABC} , we have $H_{\max}^\varepsilon(A|B)_\rho = -H_{\min}^\varepsilon(A'|C)_{\hat{\rho}}$. The state $\hat{\rho}$ has a purification on B' (in fact, this defines the embedding) that is ε -close $\rho_{A'B'C}$. Hence, the result follows. \square

5.3 Properties of the Smooth Entropies

5.3.1 Invariance under Isometries

The ε -smooth min- and max-entropies are independent of the Hilbert spaces used to represent the density operator locally, as the following lemma shows.

Proposition 5.3. *Let $\varepsilon \geq 0$ and $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Then, for all embeddings $\mathcal{H}_A \rightarrow \mathcal{H}_{A'}$ and $\mathcal{H}_B \rightarrow \mathcal{H}_{B'}$, the embedded state $\rho_{A'B'}$ satisfies*

$$H_{\min}^\varepsilon(A|B)_\rho = H_{\min}^\varepsilon(A'|B')_\rho \quad \text{and} \quad H_{\max}^\varepsilon(A|B)_\rho = H_{\max}^\varepsilon(A'|B')_\rho.$$

Proof. Let $U : \mathcal{H}_A \rightarrow \mathcal{H}_{A'}$ and $V : \mathcal{H}_B \rightarrow \mathcal{H}_{B'}$ be above-mentioned embeddings and let ρ_{ABC} be a purification of ρ . We also introduce the state $\rho_{A'B'C} = (U \otimes V)\rho_{ABC}(U^\dagger \otimes V^\dagger)$, which purifies $\rho_{A'B'}$.

We first consider the special case $\varepsilon = 0$. The property for the min-entropy then follows directly from the invariance of Φ , i.e. $\Phi_{A|B}(\rho_{AB}) = \Phi_{A'|B'}(\rho_{A'B'})$. For the max-entropy, we use $\Phi_{A|C}(\rho_{AC}) = \Phi_{A'|C}(\rho_{A'C})$, which establishes the equivalence of the max-entropies expressed in terms of the purifications ρ_{ABC} and $\rho_{A'B'C}$, respectively.

To extend this to $\varepsilon > 0$, we apply the following argument separately to the statement for the smooth min-entropy and the statement for the smooth max-entropy. The argument for the min-entropy goes as follows. We first introduce states $\hat{\rho}_{AB} \in \mathcal{B}^\varepsilon(\text{supp } \{\rho_A \otimes \rho_B\}; \rho_{AB})$ and $\tilde{\rho}_{A'B'} \in \mathcal{B}^\varepsilon(\text{supp } \{\rho_{A'} \otimes \rho_{B'}\}; \rho_{A'B'})$ that maximize $H_{\min}^\varepsilon(A|B)$ and $H_{\min}^\varepsilon(A'|B')$, respectively (cf. Proposition 5.1). Since the purified distance is non-increasing under trace non-increasing CPMs (cf. Theorem 3.4), we immediately find that

$$\hat{\rho}_{A'B'} = (U \otimes V)\hat{\rho}_{AB}(U^\dagger \otimes V^\dagger) \in \mathcal{B}^\varepsilon(\mathcal{H}_{A'B'}; \rho_{A'B'})$$

5.3 Properties of the Smooth Entropies

is a candidate for the optimization of $H_{\min}^\varepsilon(A'|B')$. Hence,

$$H_{\min}^\varepsilon(A'|B')_\rho \geq H_{\min}(A'|B')_{\hat{\rho}} = H_{\min}(A|B)_{\hat{\rho}} = H_{\min}^\varepsilon(A|B)_\rho.$$

The same argument applies in the converse direction. There, we define $\bar{\rho}_{AB}$ as the pre-image of $\bar{\rho}_{A'B'}$ under the isometry $U \otimes V$. (This is possible since $\bar{\rho}_{A'B'}$ lies in the support of $\rho_{A'} \otimes \rho_{B'}$ and, thus, in the image of $U \otimes V$.) This establishes

$$H_{\min}^\varepsilon(A|B)_\rho \geq H_{\min}(A|B)_{\bar{\rho}} = H_{\min}(A'|B')_{\bar{\rho}} = H_{\min}^\varepsilon(A|B)_\rho.$$

Therefore, equality holds. \square

Note that the above extension of a non-smooth argument to a smooth argument is quite generic and only relies on the monotonicity of the purified distance under trace non-increasing maps. We will often use variations of the above technique to lift proofs for $\varepsilon = 0$ to $\varepsilon > 0$.

5.3.2 Duality of Smooth Entropies

The *duality* relation in Lemma 4.1 extends to smooth entropies.

Theorem 5.4. *Let $\varepsilon \geq 0$ and let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ be pure. Then,*

$$H_{\max}^\varepsilon(A|B)_\rho = -H_{\min}^\varepsilon(A|C)_\rho.$$

Proof. According to Proposition 5.3, the smooth entropies are invariant under embeddings. Let $\rho_{ABC'}$ be an embedding of ρ_{ABC} into a space $\mathcal{H}_{AB} \otimes \mathcal{H}_{C'}$ with $\dim \{\mathcal{H}_{C'}\} = \max\{d_C, d_{AB}\}$. Then,

$$\begin{aligned} H_{\max}^\varepsilon(A|B)_\rho &= \min_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{AB})} H_{\max}(A|B)_{\tilde{\rho}} = \min_{\tilde{\rho} \in \mathcal{B}_p^\varepsilon(\rho_{ABC'})} H_{\max}(A|B)_{\tilde{\rho}} \\ &= - \max_{\tilde{\rho} \in \mathcal{B}_p^\varepsilon(\rho_{ABC'})} H_{\min}(A|C')_{\tilde{\rho}} \geq - \max_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{AC'})} H_{\min}(A|C')_{\tilde{\rho}} \\ &= -H_{\min}^\varepsilon(A|C')_\rho = -H_{\min}^\varepsilon(A|C)_\rho. \end{aligned}$$

We used that $\mathcal{H}_{C'}$ is chosen large enough to accommodate all purifications of states $\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{AB})$ (cf. Theorem 3.5). However, this also means that not all purifications of states $\tilde{\rho}_{AC'} \in \mathcal{B}^\varepsilon(\rho_{AC'})$ can be found in $\mathcal{B}_p^\varepsilon(\rho_{ABC'})$, which leads to the inequality.

On the other hand, we may consider the embedding $\rho_{AB'C}$ into a space $\mathcal{H}_{AC} \otimes \mathcal{H}_{B'}$ with $\dim \{\mathcal{H}_{B'}\} = \max\{d_B, d_{AC}\}$. Then, using the same arguments as above,

$$\begin{aligned} H_{\max}^\varepsilon(A|B)_\rho &= H_{\max}^\varepsilon(A|B')_\rho = \min_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{AB'})} H_{\max}(A|B')_{\tilde{\rho}} \\ &\leq \min_{\tilde{\rho} \in \mathcal{B}_p^\varepsilon(\rho_{AB'C})} H_{\max}(A|B')_{\tilde{\rho}} = - \max_{\tilde{\rho} \in \mathcal{B}_p^\varepsilon(\rho_{AB'C})} H_{\min}(A|C)_{\tilde{\rho}} \\ &= - \max_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{AC})} H_{\min}(A|C)_{\tilde{\rho}} = -H_{\min}^\varepsilon(A|C)_\rho. \end{aligned} \quad \square$$

5. SMOOTH ENTROPIES

5.3.3 Relation between Smooth Entropies

We have established in Proposition 4.3 that the min-entropy cannot exceed the max-entropy for normalized states. Here, this result is extended to smooth entropies. (See also [VDTR11] for an alternative proof, which inspired the one provided here.)

Proposition 5.5. *Let $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$ and $\varepsilon, \varepsilon' \geq 0$ s.t. $\varepsilon + \varepsilon' < 1$. Then,*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq H_{\max}^{\varepsilon'}(A|B)_{\rho} + \log \frac{1}{1 - (\varepsilon + \varepsilon')^2}.$$

Proof. First, note that we can always embed \mathcal{H}_A into a larger space $\mathcal{H}_{A'}$ such that there exists a normalized state $\tilde{\rho}_{A'B} \in \mathcal{B}^{\varepsilon}(\rho_{A'B})$ that maximizes $H_{\min}^{\varepsilon}(A'|B) = H_{\min}^{\varepsilon}(A|B)$ (cf. Lemma 5.2). Similarly, by duality, there exists a normalized state $\bar{\rho}_{A'B'} \in \mathcal{B}^{\varepsilon'}(\rho_{A'B'})$ that minimizes $H_{\max}^{\varepsilon'}(A'|B') = H_{\max}^{\varepsilon'}(A|B)$. The system B' contains B and the purification of the additional weight introduced to normalize the dual smooth min-entropy.

Hence, by definition of the smooth min-entropy, there exists a state $\sigma \in \mathcal{S}_=(\mathcal{H}_{B'})$ such that $\bar{\rho}_{A'B'} \leq 2^{-\lambda} \mathbb{1}_{A'} \otimes \sigma_{B'}$ with $\lambda = H_{\min}^{\varepsilon}(A'|B')_{\rho}$. Thus,

$$\begin{aligned} H_{\max}^{\varepsilon'}(A|B)_{\rho} &= H_{\max}^{\varepsilon'}(A'|B')_{\rho} \geq \log \left\| \sqrt{\bar{\rho}_{A'B'}} \sqrt{\mathbb{1}_{A'} \otimes \sigma_{B'}} \right\|_1^2 \\ &\geq \lambda + \log \left\| \sqrt{\bar{\rho}_{A'B'}} \sqrt{\tilde{\rho}_{A'B'}} \right\|_1^2 = \lambda + \log (1 - P^2(\bar{\rho}_{A'B'}, \tilde{\rho}_{A'B'})) \\ &\geq H_{\min}^{\varepsilon}(A|B)_{\rho} - \log \frac{1}{1 - (\varepsilon + \varepsilon')^2}. \end{aligned}$$

The first inequality follows from the definition of the smooth max-entropy, together with the fact that we took a particular σ instead of optimizing over all σ . The second inequality is a simple application of the operator inequality that defines $H_{\min}^{\varepsilon}(A'|B)$. Finally, the triangle inequality for the purified distance establishes $P(\bar{\rho}, \tilde{\rho}) \leq \varepsilon + \varepsilon'$. \square

Remark 5.6. *The term $-\log(1 - (\varepsilon + \varepsilon')^2)$ can be reduced by the use of Eq. (3.8) instead of the triangle inequality for the purified distance. Hence,*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq H_{\max}^{\varepsilon'}(A|B)_{\rho} + 2 \log \frac{1}{1 - \left(\varepsilon \sqrt{1 - \varepsilon'^2} + \varepsilon' \sqrt{1 - \varepsilon^2} \right)^2}.$$

The range of allowed pairs $\{\varepsilon, \varepsilon'\}$ is extended to those satisfying $\arcsin(\varepsilon) + \arcsin(\varepsilon') < \frac{\pi}{2}$. In particular, this means that the term is finite if we choose $\varepsilon' = 1 - \varepsilon$, for any $0 < \varepsilon < 1$. (See also Figure 3.1.)

Proposition 5.5 implies that smoothing states that have similar min- and max-entropies has almost no effect. More precisely, let $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$ be such that $H_{\min}(A|B)_{\rho} = H_{\max}(A|B)_{\rho} - \delta$. Then,

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq H_{\max}(A|B)_{\rho} - \log(1 - \varepsilon^2) = H_{\min}(A|B)_{\rho} + \delta - \log(1 - \varepsilon^2).$$

For $\delta = 0$, this inequality is tight. Note that the smoothed state $\tilde{\rho} = \rho(1 - \varepsilon^2) \approx_{\varepsilon} \rho$ reaches equality in this case. A corresponding inequality can be derived for the smooth max-entropy.

5.4 Data Processing Inequalities

We expect measures of uncertainty about the system A given side information B to be non-decreasing under local physical operations (e.g. measurements or unitary evolutions) applied to the B system. Such operations can be described most generally by *TP-CPMs*. Here, we show that the smooth entropies, $H_{\min}^{\varepsilon}(A|B)$ and $H_{\max}^{\varepsilon}(A|B)$, have this property.

Another data processing inequality concerns rank-1 *projective measurements* of the system A . Such measurements can be described in terms of an orthonormal basis $\{|x\rangle\}$ of \mathcal{H}_A and a measurement TP-CPM $\mathcal{M} \in \mathcal{T}(\mathcal{H}_A, \mathcal{H}_X)$ from \mathcal{H}_A to $\mathcal{H}_X \cong \mathcal{H}_A$, which maps ρ to $\sum_x \langle x|\rho|x\rangle |x\rangle\langle x|$. We expect that the uncertainty about the system A as well as the smooth entropies, $H_{\min}^{\varepsilon}(A|B)$ and $H_{\max}^{\varepsilon}(A|B)$, do not decrease when such a measurement is executed on the A system.

In fact, we show a more general theorem that encompasses the two example data processing properties above. (Note, in particular, that \mathcal{M} is a *unital* map.)

Theorem 5.7 (Generalized Data Processing). *Let $\varepsilon \geq 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$. Moreover, let \mathcal{E} be a sub-unital and trace non-increasing CPM from \mathcal{H}_A to \mathcal{H}'_A , let \mathcal{F} be a trace non-increasing CPM from \mathcal{H}_B to $\mathcal{H}_{B'}$ and let \mathcal{G} be a trace non-increasing CPM from \mathcal{H}_C to $\mathcal{H}_{C'}$. Then, the state $\tau_{A'B'C'} = (\mathcal{G} \circ \mathcal{F} \circ \mathcal{E})[\rho_{ABC}]$ satisfies*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq H_{\min}^{\varepsilon}(A'|B')_{\tau}.$$

Furthermore, if \mathcal{E} , \mathcal{F} and \mathcal{G} are also trace preserving, then

$$H_{\max}^{\varepsilon}(A|B)_{\rho} \leq H_{\max}^{\varepsilon}(A'|B')_{\tau}.$$

Proof. We first prove the result for the smooth min-entropy. Let $\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho_{AB})$ be the state that maximizes the smooth min-entropy. Then, data processing of Φ (Properties vi. and vii.) implies

$$\begin{aligned} H_{\min}^{\varepsilon}(A|B)_{\rho} &= H_{\min}(A|B)_{\tilde{\rho}} \\ &= -\log \Phi_{A|B}(\tilde{\rho}) \\ &\leq -\log \Phi_{A'|B'}((\mathcal{F} \circ \mathcal{E})[\tilde{\rho}]) \\ &\leq -\log \Phi_{A'|B'}((\mathcal{G} \circ \mathcal{F} \circ \mathcal{E})[\tilde{\rho}]). \end{aligned}$$

We now introduce the state $\tilde{\tau} = (\mathcal{G} \circ \mathcal{F} \circ \mathcal{E})[\tilde{\rho}]$, which is ε -close to τ due to the monotonicity of the purified distance under trace non-increasing maps

5. SMOOTH ENTROPIES

(cf. Theorem 3.4). Thus, we conclude

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq -\log \Phi_{A'|B'}(\tilde{\tau}) = H_{\min}(A'|B')_{\tilde{\tau}} \leq H_{\min}^{\varepsilon}(A'|B')_{\tau}.$$

To prove the result for the max-entropy, we take advantage of the *Stinespring dilation* of two TP-CPM maps \mathcal{E} and \mathcal{F} . Namely, we introduce the isometries $U : A \rightarrow A'A''$ and $V : B \rightarrow B'B''$ and the state $\tau_{A'A''B'B''} = (U \otimes V)\rho_{AB}(U^{\dagger} \otimes V^{\dagger})$ of which $\tau_{A'B'}$ is a marginal. Let $\tilde{\tau} \in \mathcal{B}^{\varepsilon}(\tau_{A'A''B'B''})$ be the state that minimizes the smooth max-entropy $H_{\max}^{\varepsilon}(A'|B')_{\tau}$. Then,

$$\begin{aligned} H_{\max}^{\varepsilon}(A'|B')_{\tau} &= \max_{\sigma} \log \left\| \sqrt{\tilde{\tau}_{A'B'}} \sqrt{\mathbb{1}_{A'} \otimes \sigma_{B'}} \right\|_1^2 \\ &\geq \max_{\sigma} \log \left\| \sqrt{\tilde{\tau}_{A'B'}} \sqrt{\text{tr}_{A''}(\Pi_{A'A''}) \otimes \sigma_{B'}} \right\|_1^2. \end{aligned} \quad (5.2)$$

We introduced the projector $\Pi_{A'A''} = UU^{\dagger}$ onto the image of U , which exhibits the following property due to the fact that \mathcal{E} is sub-unital:

$$\text{tr}_{A''}(\Pi) = \text{tr}_{A''}(U\mathbb{1}_A U^{\dagger}) = \mathcal{E}[\mathbb{1}_A] \leq \mathbb{1}_{A'}.$$

The inequality in (5.2) is then a result of the fact that the fidelity is non-increasing when an argument A is replaced by a smaller argument $B \leq A$. (Property v. of the fidelity). Next, we use the monotonicity of the fidelity under partial trace (3.2) to bound (5.2) further.

$$\begin{aligned} H_{\max}^{\varepsilon}(A'|B')_{\tau} &\geq \max_{\sigma} \log \left\| \sqrt{\tilde{\tau}_{A'A''B'B''}} \sqrt{\Pi_{A'A''} \otimes \sigma_{B'B''}} \right\|_1^2 \\ &= \max_{\sigma} \log \left\| \sqrt{\Pi_{A'A''} \tilde{\tau}_{A'A''B'B''} \Pi_{A'A''}} \sqrt{\mathbb{1}_{A'A''} \otimes \sigma_{B'B''}} \right\|_1^2 \\ &= H_{\max}(A'A''|B'B'')_{\tilde{\tau}}. \end{aligned}$$

Finally, we note that the projector $\hat{\tau} = \Pi \tilde{\tau} \Pi \in \mathcal{B}^{\varepsilon}(\rho_{A'A''B'B''})$ due to the monotonicity of the purified distance under trace non-increasing maps. Hence, we established $H_{\max}^{\varepsilon}(A'|B')_{\tau} \geq H_{\max}^{\varepsilon}(A'A''|B'B'')_{\tau} = H_{\max}^{\varepsilon}(A|B)_{\rho}$, where the last equality follows due to the invariance of the max-entropy under embeddings (Proposition 5.3). \square

Note that a generalization of the data processing results to trace non-increasing maps is not possible for the max-entropy. For example, the max-entropy can decrease when a projection is applied to the B system.

5.5 Classical Information

Here we discuss smooth entropies of quantum states that encode partially classical information. To maintain full generality in the following arguments, we consider states on a four-partite system composed of X , Y , A and B , where X and Y are classical *registers*. The smooth entropies are

evaluated for the classical register X and the quantum register A conditioned on classical side information Y and quantum side information B .

In the following, we say that the state $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{XYAB})$ is *classical-classical-quantum-quantum (CCQQ)* if it can be written in the form

$$\rho_{XYAB} = \sum_{x,y} p_{xy} |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \otimes \omega_{AB}^{xy}, \quad \text{where } \omega_{AB}^{xy} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB}), \quad (5.3)$$

the kets $\{|x\rangle\}$, $\{|y\rangle\}$ are orthonormal bases of \mathcal{H}_X and \mathcal{H}_Y , respectively, and $\{p_{xy}\}$ is a probability distribution. This state has the property that it remains invariant under a projective measurement of X in the basis $\{|x\rangle\}$ and under a projective measurement of Y in the basis $\{|y\rangle\}$. Formally, we define measurement *TP-CPMs*

$$\mathcal{M}_X : \rho \mapsto \sum_x \langle x|\rho|x\rangle_X |x\rangle\langle x|_X \quad \text{and} \quad \mathcal{M}_Y : \rho \mapsto \sum_y \langle y|\rho|y\rangle_Y |y\rangle\langle y|_Y \quad (5.4)$$

and note that $\mathcal{M}_X[\rho] = \mathcal{M}_Y[\rho] = \rho$. Furthermore, given an arbitrary state $\tau \in \mathcal{S}_{\leq}(\mathcal{H}_{XYAB})$, it is easy to verify that the measured state $(\mathcal{M}_Y \circ \mathcal{M}_X)[\tau]$ is of CCQQ form (5.3).

Partially classical states of the form (5.3) can be purified in a way that preserves their structure. For this, we introduce purifying Hilbert spaces $\mathcal{H}_{X'} \cong \mathcal{H}_X$, $\mathcal{H}_{Y'} \cong \mathcal{H}_Y$ and $\mathcal{H}_C \cong \mathcal{H}_{AB}$. Then, one may purify the states ω_{AB}^{xy} individually on $\mathcal{S}_{\leq}(\mathcal{H}_{ABC})$, which allows us to write down a purification $|\rho\rangle \in \mathcal{H}_{XX'YY'ABC}$ of the form

$$|\rho\rangle_{XX'YY'ABC} = \sum_{x,y} \sqrt{p_{xy}} |x\rangle_X \otimes |x\rangle_{X'} \otimes |y\rangle_Y \otimes |y\rangle_{Y'} \otimes |\omega^{xy}\rangle_{ABC}. \quad (5.5)$$

It is easy to verify that purifications of this form commute with the following projectors,

$$\Pi_{XX'} := \sum_x |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \quad \text{and} \quad \Pi_{YY'} := \sum_y |y\rangle\langle y|_Y \otimes |y\rangle\langle y|_{Y'}. \quad (5.6)$$

In the converse, any state $\tau \in \mathcal{S}_{\leq}(\mathcal{H}_{XX'YY'AB})$ that commutes with both $\Pi_{XX'}$ and $\Pi_{YY'}$ has a CCQQ marginal τ_{XYAB} . We say that the two systems X and X' as well as Y and Y' are *coherent classical* pairs.

Some important properties of min- and max-entropies of coherent classical states are discussed in Appendix A, Section A.2 and will be used in the following.

5.5.1 Smoothing of Classical States

First, we show that—in order to smooth min- and max-entropies—it is sufficient to consider a ball of close states that are classical on the same subsystems as the original state. This is formalized in the following proposition.

5. SMOOTH ENTROPIES

Proposition 5.8 (Classical Smoothing). *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{XYAB})$ be classical on X and Y and let $\varepsilon \geq 0$. Then, there exist states $\bar{\rho}, \tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho)$ that are classical on X and Y such that*

$$\begin{aligned} H_{\min}^{\varepsilon}(XA|YB)_{\rho} &= H_{\min}(XA|YB)_{\bar{\rho}} \quad \text{and} \\ H_{\max}^{\varepsilon}(XA|YB)_{\rho} &= H_{\max}(XA|YB)_{\tilde{\rho}}. \end{aligned}$$

Proof. We first prove the statement for the min-entropy. By definition of the smooth min-entropy, there exists a state $\hat{\rho} \in \mathcal{B}^{\varepsilon}(\rho)$ such that $H_{\min}^{\varepsilon}(XA|YB)_{\rho} = H_{\min}(XA|YB)_{\hat{\rho}}$. We propose the CCQQ state $\bar{\rho} = (\mathcal{M}_X \circ \mathcal{M}_Y)[\hat{\rho}]$ using the measurement in (5.4) as a candidate. Due to the data processing inequalities in Theorem 5.7, we have

$$H_{\min}(XA|YB)_{\hat{\rho}} \leq H_{\min}(XA|YB)_{\bar{\rho}}.$$

Furthermore, $\bar{\rho} \approx_{\varepsilon} \rho$ due to Theorem 3.4 and, hence, $\bar{\rho}$ satisfies all conditions of the proposition.

To prove the statement for the max-entropy, we consider the dual problem for the min-entropy using the purification $|\rho\rangle$ in (5.5). Namely, we need to show that there exists a state $\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho)$ that satisfies

$$H_{\min}^{\varepsilon}(XA|X'Y'C)_{\rho} = H_{\min}(XA|X'Y'C)_{\tilde{\rho}},$$

commutes with $\Pi_{XX'}$, and is classical on Y' . Due to the first part of this lemma, there exists a state $\hat{\rho} \in \mathcal{B}^{\varepsilon}(\rho)$ such that $H_{\min}^{\varepsilon}(XA|X'Y'C)_{\rho} = H_{\min}(XA|X'Y'C)_{\hat{\rho}}$ and $\hat{\rho}$ is classical on Y' . We propose the state $\tilde{\rho} = \Pi_{XX'}\hat{\rho}\Pi_{XX'}$ as a candidate. Then, Lemma A.3 and Theorem 3.4 establish that $\tilde{\rho}$ satisfies all conditions, which conclude the proof. \square

5.5.2 Entropy of Classical Information

Here, we bound the smooth entropies of a classical system X . For full generality, we consider the state ρ_{XAB} , which is simply the state (5.3) with trivial classical subsystem Y .

Proposition 5.9. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{XAB})$ be classical on X and let $0 \leq \varepsilon < 1$. Then,*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \leq H_{\min}^{\varepsilon}(XA|B)_{\rho} \leq H_{\min}^{\varepsilon}(A|B)_{\rho} + \log d_X \quad \text{and} \quad (5.7)$$

$$H_{\max}^{\varepsilon}(A|B)_{\rho} \leq H_{\max}^{\varepsilon}(XA|B)_{\rho} \leq H_{\max}^{\varepsilon}(A|B)_{\rho} + \log d_X. \quad (5.8)$$

Proof. To show the first inequality of (5.7), we consider the dual problem for the max-entropy. Namely, we need to show that

$$H_{\max}^{\varepsilon}(XA|X'C)_{\rho} \leq H_{\max}^{\varepsilon}(A|CXX')_{\rho}$$

for the state $\rho_{XX'AC}$ that is coherent classical between X and X' . To do this, note that there exists a $\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)$ with $H_{\max}^\varepsilon(A|CXX')_\rho = H_{\max}(A|CXX')_{\tilde{\rho}}$ that has support on $\text{supp}\{\rho_A\} \otimes \text{supp}\{\rho_{XX'C}\}$ (cf. Proposition 5.1) and, thus, commutes with $\Pi_{XX'}$. The first inequality now follows after we apply Lemma A.4 to $\tilde{\rho}$ and realize that

$$H_{\max}^\varepsilon(XA|X'C)_\rho \leq H_{\max}(XA|X'C)_{\tilde{\rho}} \leq H_{\max}(A|XX'C)_{\tilde{\rho}}.$$

Since smoothing can be restricted to states that are classical on X (cf. Proposition 5.8), there exist states $\bar{\rho} \in \mathcal{B}^\varepsilon(\rho)$ and $\sigma \in \mathcal{S}_=(\mathcal{H}_B)$ such that

$$\bar{\rho}_{XAB} = \sum_x \bar{p}_x |x\rangle\langle x|_X \otimes \bar{\tau}_{AB}^x \leq 2^{-H_{\min}^\varepsilon(XA|B)_\rho} \mathbb{1}_{XA} \otimes \sigma_B.$$

Since both sides of the inequality are block diagonal in X , this implies that

$$\begin{aligned} \forall_x : \bar{p}_x \bar{\tau}_{AB}^x &\leq 2^{-H_{\min}^\varepsilon(XA|B)_\rho} \mathbb{1}_A \otimes \sigma_B \\ \implies \bar{\rho}_{AB} &= \sum_x \bar{p}_x \bar{\tau}_{AB}^x \leq d_X 2^{-H_{\min}^\varepsilon(XA|B)_\rho} \mathbb{1}_A \otimes \sigma_B. \end{aligned}$$

This implies the second inequality of (5.7).

The first inequality of (5.8) is equivalent to the statement

$$H_{\min}^\varepsilon(A|XX'C)_\rho \geq H_{\min}^\varepsilon(XA|X'C)_\rho$$

for a state ρ that is coherent classical between X and X' . The inequality then follows directly from Lemma A.3 applied to the state $\bar{\tau} \in \mathcal{B}^\varepsilon(\rho)$ that maximizes $H_{\min}^\varepsilon(XA|X'C)_\rho$ and Theorem 3.4, which establishes that $\bar{\rho} = \Pi_{XX'} \bar{\tau} \Pi_{XX'}$ is a candidate for the maximization of $H_{\min}^\varepsilon(A|XX'C)_\rho$.

The second inequality of (5.8) is shown as follows. By the definition of the smooth max-entropy, there exists a state $\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)$ and a state $\sigma \in \mathcal{S}_=(\mathcal{H}_B)$ such that

$$\begin{aligned} H_{\max}^\varepsilon(A|B)_\rho &\geq 2 \log F(\tilde{\rho}_{AB}, \mathbb{1}_A \otimes \sigma_B) \\ &\geq 2 \log F(\tilde{\rho}_{XAB}, \frac{1}{d_X} \mathbb{1}_{XA} \otimes \sigma_B) \geq H_{\max}^\varepsilon(XA|B)_\rho - \log d_X, \end{aligned}$$

where σ maximizes $H_{\max}(XA|B)_{\tilde{\rho}}$ and $\tilde{\rho}$ minimizes $H_{\max}^\varepsilon(A|B)$. \square

5.5.3 Conditioning on Classical Information

We investigate the maximum amount of information a classical register X can contain about a quantum state A . (Various relations of this type have appeared in the literature, see, e.g. [RR10, WTHR11]. The following lemma generalizes these results.)

5. SMOOTH ENTROPIES

Proposition 5.10. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{XAB})$ be classical on Y and let $0 \leq \varepsilon < 1$. Then,*

$$H_{\min}^{\varepsilon}(A|B)_{\rho} \geq H_{\min}^{\varepsilon}(A|BY)_{\rho} \geq H_{\min}^{\varepsilon}(A|B)_{\rho} - \log d_Y \quad \text{and} \quad (5.9)$$

$$H_{\max}^{\varepsilon}(A|B)_{\rho} \geq H_{\max}^{\varepsilon}(A|BY)_{\rho} \geq H_{\max}^{\varepsilon}(A|B)_{\rho} - \log d_Y. \quad (5.10)$$

Proof. The first inequalities of both (5.9) and (5.10) directly follow from data processing (cf. Theorem 5.7) applied to the TP-CPM tr_Y .

Let $\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho_{ABY})$ be a state that achieves the maximal min-entropy for $H_{\min}^{\varepsilon}(AY|B)_{\rho} = H_{\min}(AY|B)_{\tilde{\rho}} = \lambda$. In particular, there exists a state $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ s.t.

$$\tilde{\rho}_{ABY} \leq 2^{-\lambda} \mathbb{1}_{AY} \otimes \sigma_B = d_Y 2^{-\lambda} \mathbb{1}_A \otimes \pi_Y \otimes \sigma_B.$$

This implies that $H_{\min}^{\varepsilon}(A|BY)_{\rho} \geq H_{\min}^{\varepsilon}(AY|B)_{\rho} - \log d_Y$. The second inequality in (5.9) then directly follows from Proposition 5.9.

To establish the second inequality in (5.7), we prove the dual property, i.e. for the purification $\rho_{ABCY'Y'}$ that is coherent classical between Y and Y' , we show that $H_{\min}^{\varepsilon}(A|CY')_{\rho} \leq H_{\min}^{\varepsilon}(A|CYY')_{\rho} + \log d_Y$. Let $\tilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho_{ACY'})$ and $\sigma \in \mathcal{S}_{\leq}(\mathcal{H}_{CY'})$ be such that

$$\tilde{\rho}_{ACY'} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_{CY'}, \quad \text{where} \quad \lambda = H_{\min}^{\varepsilon}(A|CY')_{\rho}.$$

Lemma A.2 implies that any extension $\tilde{\rho}_{ACY'Y'} \in \mathcal{B}^{\varepsilon}(\rho)$ satisfies

$$\tilde{\rho}_{ACY'Y'} \leq d_Y \tilde{\rho}_{ACY'} \otimes \mathbb{1}_{Y'} \leq 2^{-\lambda} d_Y \mathbb{1}_{YA} \otimes \sigma_{CY'}.$$

Hence, we find the following inequality

$$H_{\min}(AY|CY')_{\tilde{\rho}} \geq H_{\min}^{\varepsilon}(A|CY')_{\rho} - \log d_Y.$$

An application of Lemma A.3 to the min-entropy on the lhs. establishes

$$H_{\min}(A|CYY')_{\tilde{\tau}} \geq H_{\min}^{\varepsilon}(A|CY')_{\rho} - \log d_Y, \quad \text{where} \quad \tilde{\tau} = \Pi_{Y'Y'} \rho \Pi_{Y'Y'}$$

is coherent classical between Y and Y' . Since ρ has the same property, Theorem 3.4 implies that $\tilde{\tau} \in \mathcal{B}^{\varepsilon}(\rho)$, which concludes the proof. \square

5.5.4 Functions on Classical Registers

Let us again consider the state ρ_{AXBY} of (5.3). On the one hand, applying a function $f : Y \rightarrow Z$ on the register Y is a special case of data processing (cf. Theorem 5.7) and we immediately find

$$H_{\min}^{\varepsilon}(A|BY)_{\rho} \leq H_{\min}^{\varepsilon}(A|BZ)_{\rho} \quad \text{and} \quad H_{\max}^{\varepsilon}(A|BY)_{\rho} \leq H_{\max}^{\varepsilon}(A|BZ)_{\rho}.$$

On the other hand, Proposition 5.10 can be used to show that applying a function on the register X cannot increase the smooth entropies. (Note that for the min-entropy, using its interpretation as a *guessing probability*, this corresponds to the intuitive statement that it is always at least as hard to guess the input of a function than to guess its output.)

Proposition 5.11. *Let $\rho_{XAB} = \sum p_x |x\rangle\langle x| \otimes \omega_{AB}^x$ with $\omega_{AB}^x \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ be classical on X . Furthermore, let $0 \leq \varepsilon < 1$ and let $f : X \rightarrow Z$ be a function. Then, the state $\tau_{ZAB} = \sum p_x |f(x)\rangle\langle f(x)|_Z \otimes \omega_{AB}^x$ satisfies*

$$H_{\min}^{\varepsilon}(ZA|B)_{\tau} \leq H_{\min}^{\varepsilon}(XA|B)_{\rho} \quad \text{and} \quad H_{\max}^{\varepsilon}(ZA|B)_{\tau} \leq H_{\max}^{\varepsilon}(XA|B)_{\rho}.$$

Proof. A possible *Stinespring dilation* of f is given by the isometry $U : |x\rangle_X \mapsto |x\rangle_{X'} \otimes |f(x)\rangle_Z$ followed by a partial trace over X' . Applying U on ρ_{XAB} , we get

$$\tau_{XZAB} := U\rho_{XAB}U^{\dagger} = \sum p_x |x\rangle\langle x|_{X'} \otimes |f(x)\rangle\langle f(x)|_Z \otimes \omega_{AB}^x$$

which is classical on X' and Z and an extension of τ_{ZAB} . Hence, the invariance under isometries of the smooth entropies (cf. Proposition 5.3) in conjunction with Proposition 5.10 implies

$$H_{\min}^{\varepsilon}(XA|B)_{\rho} = H_{\min}^{\varepsilon}(XZA|B)_{\tau} \geq H_{\min}^{\varepsilon}(ZA|B)_{\tau}.$$

An equal argument for the smooth max-entropy concludes the proof. \square

5.6 Chain Rules

The *chain rule* for the von Neumann (or Shannon) entropy is the relation $H(AB|C) = H(A|BC) + H(B|C)$. To see that the von Neumann entropy satisfies this relation, simply substitute the definition of the conditional entropy, i.e. $H(A|B) = H(AB) - H(B)$. For the smooth min- and max-entropies, this rule does not hold with equality.

Instead, we give a collection of inequalities that replace the chain rule for the smooth min- and max-entropies. These were recently derived in [Vit11, VDTR11]. (See also [TSSR11, BCC⁺10] for preliminary results.) These chain rules generalize some of the above results for classical registers to general quantum systems; however, they introduce an additional, necessary smoothing parameter, ε , and correction terms in $\log(2/\varepsilon^2)$ that do not appear in the results of the previous section. This is due to the fact that, previously, we gave bounds in terms of the dimension of the classical register instead of its max-entropy. This simplifies the analysis and is sufficient in many applications involving classical information.

The chain rules provided here should thus be considered complementary to the other results of this chapter.

Theorem 5.12. *Let $\varepsilon > 0$, $\varepsilon' \geq 0$, $\varepsilon'' \geq 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$. Then, the*

5. SMOOTH ENTROPIES

following chain rules hold

$$H_{\min}^{\varepsilon+2\varepsilon'+\varepsilon''}(AB|C)_\rho \geq H_{\min}^{\varepsilon'}(A|BC)_\rho + H_{\min}^{\varepsilon''}(B|C)_\rho - \log \frac{2}{\varepsilon^2}, \quad (5.11)$$

$$H_{\max}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB|C)_\rho \leq H_{\max}^{\varepsilon'}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + \log \frac{2}{\varepsilon^2}, \quad (5.12)$$

$$H_{\min}^{\varepsilon+3\varepsilon'+2\varepsilon''}(A|BC)_\rho \geq H_{\min}^{\varepsilon'}(AB|C)_\rho - H_{\max}^{\varepsilon''}(B|C)_\rho - 2 \log \frac{2}{\varepsilon^2}, \quad (5.13)$$

$$H_{\max}^{2\varepsilon+\varepsilon'+2\varepsilon''}(A|BC)_\rho \leq H_{\max}^{\varepsilon'}(AB|C)_\rho - H_{\min}^{\varepsilon''}(B|C)_\rho + 3 \log \frac{2}{\varepsilon^2}, \quad (5.14)$$

$$H_{\min}^{2\varepsilon+\varepsilon'+2\varepsilon''}(B|C)_\rho \geq H_{\min}^{\varepsilon'}(AB|C)_\rho - H_{\max}^{\varepsilon''}(A|BC)_\rho - 3 \log \frac{2}{\varepsilon^2}, \quad (5.15)$$

$$H_{\max}^{\varepsilon+3\varepsilon'+2\varepsilon''}(B|C)_\rho \leq H_{\max}^{\varepsilon'}(AB|C)_\rho - H_{\min}^{\varepsilon''}(A|BC)_\rho + 2 \log \frac{2}{\varepsilon^2}. \quad (5.16)$$

For a proof, we refer to [Vit11]. Note that Eqs. (5.12), (5.14) and (5.16) follow by duality of the smooth entropies from Eqs. (5.11), (5.15) and (5.13), respectively.

Chapter 6

The Quantum Asymptotic Equipartition Property

The classical asymptotic equipartition property (AEP) is the statement that, in the limit of a large number of identical repetitions of a random experiment, the output sequence is virtually certain to come from the typical set, each member of which is almost equally likely. In this chapter, expanding on previous results in [TCR09], a fully quantum generalization of this property is shown, where both the output of the experiment and side information are quantum. We give an explicit bound on the convergence, which is independent of the dimensionality of the side information.

6.1 Introduction and Related Work

One of the pivotal results of classical information theory is the *asymptotic equipartition property* (AEP). This result justifies the use of the Shannon entropy to characterize many information theoretic tasks. It concerns a sequence of *independent and identically distributed* (i.i.d.) random variables $X^n = (X_1, X_2, \dots, X_n)$ and states that [CT91]

$$-\frac{1}{n} \log P_{X^n}(x^n) \rightarrow H(X)_P \quad \text{in probability .}$$

Here, $P_{X^n}(x^n) = P_{X^n}(x_1, x_2, \dots, x_n) = P_X(x_1)P_X(x_2) \dots P_X(x_n)$ denotes the probability of the independent events x_1, x_2, \dots, x_n . More precisely, the AEP states that, for any $0 < \varepsilon < 1$, any $\mu > 0$ and for large enough n , a randomly chosen sequence (x_1, x_2, \dots, x_n) is with probability more than $1 - \varepsilon$ in a *typical set* of sequences that satisfy

$$H(X) - \mu < -\frac{1}{n} \log P_{X^n}(x^n) < H(X) + \mu . \quad (6.1)$$

This typical set of sequences of length n is denoted A_μ^n .

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

The AEP can be stated equivalently in terms of entropies. Recall the min- and max-entropies of a classical random variable X^n , which satisfy

$$\begin{aligned} H_{\min}(X^n)_P &= H_{\infty}(X^n)_P = \min_{x^n} -\log P_{X^n}(x^n) \quad \text{and} \\ H_{\max}(X^n)_P &= H_{1/2}(X^n)_P \leq \max_{x^n, P(x^n) \neq 0} -\log P_{X^n}(x^n). \end{aligned} \quad (6.2)$$

The bounds in (6.1) can now be expressed in terms of the smooth min- and max-entropy. For this purpose, we consider the probability distribution Q which restricts P to the typical set. It is formally defined as

$$Q_{X^n}(x^n) = \begin{cases} P_{X^n}(x^n)/c & \text{if } x^n \in A_{\mu}^n, \\ 0 & \text{else} \end{cases}, \quad \text{where } c = \sum_{x^n \in A_{\mu}^n} P_{X^n}(x^n).$$

The AEP now states that for large enough n , the fidelity between the distributions P and Q satisfies

$$F(P, Q)^2 = \left(\sum_{x^n} \sqrt{P_{X^n}(x^n) Q_{X^n}(x^n)} \right)^2 = \sum_{x^n \in A_{\mu}^n} P_{X^n}(x^n) > 1 - \varepsilon,$$

and, thus, the purified distance is upper bounded by $\sqrt{\varepsilon}$. Finally, the AEP in (6.1), in conjunction with (6.2), implies that, for any $0 < \varepsilon < 1$, $\mu > 0$ and large enough n ,

$$\begin{aligned} \frac{1}{n} H_{\min}^{\sqrt{\varepsilon}}(X^n)_P &> \frac{1}{n} H_{\min}(X^n)_Q > H(X)_P - \mu \quad \text{and} \\ \frac{1}{n} H_{\max}^{\sqrt{\varepsilon}}(X^n)_P &< \frac{1}{n} H_{\max}(X^n)_Q < H(X)_P + \mu. \end{aligned} \quad (6.3)$$

Conversely, its entropic form (6.3) implies an AEP of the form (6.1). We only roughly sketch this argument here. First note that the fact that the smooth min-entropy converges to the von Neumann entropy implies the existence of a set \tilde{A}^n with the property $\max\{P(x^n) : x^n \in \tilde{A}^n\} \approx 2^{-nH(X)}$ and $x^n \in \tilde{A}^n$ probability almost 1. Similarly, the convergence of the max-entropy (together with the fact that the smooth max-entropy is related to smooth Rényi entropy of order 0) implies the existence of another set \bar{A}^n with $|\bar{A}^n| \approx 2^{nH(X)}$ and $x^n \in \bar{A}^n$ with probability almost 1. Since both these sets occur with probability almost 1, it is easy to see that their intersection, $\tilde{A}^n \cap \bar{A}^n$, constitutes a typical set.

The entropic form of the AEP explains the crucial role of the *von Neumann entropy* to describe information theoretic tasks. While operational quantities in information theory (such as the amount of extractable randomness, the minimal length of compressed data and channel capacities) can naturally be expressed in terms of smooth entropies in the *one-shot* setting, the von Neumann entropy is recovered if we consider a large number of independent repetitions of the task.

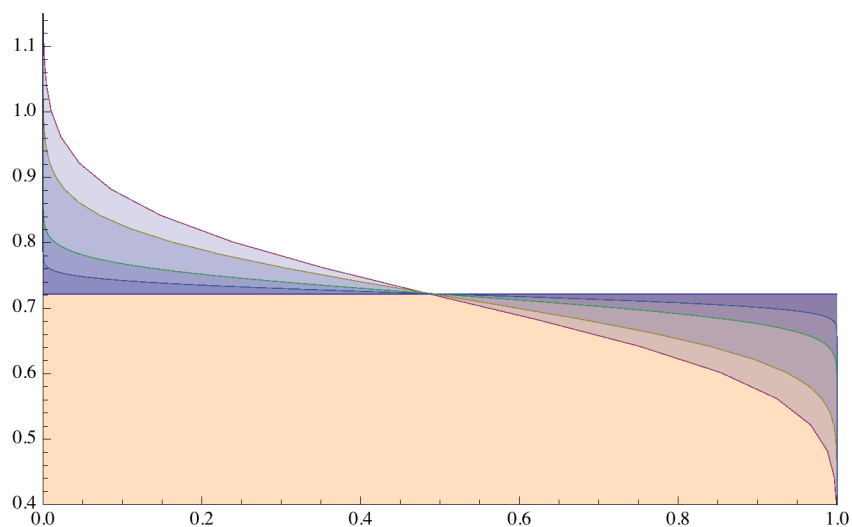


Figure 6.1: *Emergence of Typical Set*. Plot of the surprisal, $-\log P(x^n)$, over the cumulated probability of the events. We consider n independent Bernoulli trials with $p = 0.2$. The curves for $n = \{50, 100, 500, 2500\}$ converge to the typical set (indicated by the horizontal line). The min-entropy, $H_{\min}(X^n)$, is the minimum of the corresponding curve, the max-entropy, $H_{\max}(X^n)$, is upper bounded by its maximum. Moreover, the respective ε -smooth entropies can be approximately obtained by cutting off a probability ε from each side of the x -axis and taking the minima or maxima of the remaining curve.

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

Moreover, the entropic approach to asymptotic equipartition lends itself to a generalization to the quantum setting. Note that the traditional approach, which considers the AEP as a statement about (conditional) probabilities, does not have a natural quantum generalization due to the fact that we do not know a suitable generalization of conditional probabilities to quantum side information.

Here, we want to show that the smooth conditional min- and max-entropies converge to the von Neumann entropy in the *i.i.d. limit*. Recall that the conditional von Neumann entropy of a state ρ is defined as

$$H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B), \quad \text{where} \quad H(\rho) = -\text{tr}(\rho \log \rho). \quad (6.4)$$

This convergence can be shown in several ways, for example through the use of chain rules for smooth entropies [BR11], which reduce the problem to the classical AEP, or through the use of typical subspaces (cf. e.g. [NC00]). While these techniques achieve the desired asymptotic limit, they fail to give good bounds on the convergence for finite n . More precisely, we are interested in the difference between the smooth entropies and the von Neumann entropy for finite n . This distance is in general a function of n , ε and some properties of the quantum state under consideration.

We call such a relation a fully quantum AEP because both the A and B systems are general quantum systems.

For the smooth conditional min- and max-entropy, the convergence for finite n was first analyzed by Holenstein and Renner [HR11] for classical probability distributions. Renner also generalized these arguments to the quantum setting [Ren05]. He shows that¹, for any $\varepsilon > 0$ and for tensor product states $\rho_{AB}^{\otimes n} = \rho_{AB} \otimes \rho_{AB} \otimes \dots \otimes \rho_{AB}$, it holds that

$$\frac{1}{n} H_{\min}^\varepsilon(A^n|B^n)_\rho \geq H(A|B)_\rho - \frac{\delta}{\sqrt{n}}, \quad (6.5)$$

where δ scales with the product of $H_0(A)_\rho$ and $\sqrt{\log(1/\varepsilon)}$.²

On a technical level, our results are related to recent findings in quantum hypothesis testing [AMAV07, ANSV08, NO00].

6.1.1 Main Contributions

This chapter is based on [TCR09]; however, the results presented here are strictly more general as we extended certain proofs to relative entropies and a more general class of operators. This allows us to plug the inequalities into other arguments, where (e.g. smoothed) states are not necessarily normalized.

¹Note that the smoothing of the min-entropy was defined differently in [Ren05].

²Note that $H_0(A)_\rho = d_A$ if the state ρ has full support on \mathcal{H}_A ; thus, this bound depends on the dimension of the A system.

The main result of this chapter establishes that δ in (6.5) scales with $H_{\max}(A|B)$ and $H_{\max}(A|R)$ instead of $H_0(A)$, where R is a reference system purifying ρ_{AB} . These conditional entropies measure the correlations between the subsystems A , B and R and are often much smaller than $H_0(A)$. In particular, conditional entropies do not depend on the Hilbert space dimension of any subsystem, which is useful in the context of quantum cryptography, where these dimensions are generally unknown.¹

Result 5. *For any $\varepsilon > 0$ and n large enough, we have*

$$\begin{aligned} \frac{1}{n} H_{\min}^{\varepsilon}(A^n|B^n) &\geq H(A|B) - \frac{\delta(\varepsilon, v)}{\sqrt{n}} \quad \text{and} \\ \frac{1}{n} H_{\max}^{\varepsilon}(A^n|B^n) &\leq H(A|B) + \frac{\delta(\varepsilon, v)}{\sqrt{n}}, \end{aligned}$$

where $\delta(\varepsilon, v) = 4 \log v \sqrt{\log(2/\varepsilon^2)}$ and $v = \sqrt{2^{H_{\max}(A|B)}} + \sqrt{2^{H_{\max}(A|R)}} + 1$.

Together with converse bounds, this implies the fully quantum asymptotic equipartition property.

Result 6. *The smooth entropies converge to the von Neumann entropy in the asymptotic limit of many copies. For any $0 < \varepsilon < 1$, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^{\varepsilon}(A^n|B^n) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\max}^{\varepsilon}(A^n|B^n) = H(A|B).$$

Note, in particular, that our statement holds for any $0 < \varepsilon < 1$. This improves on to the preliminary results in [TCR09], where the converse bound was only shown in the limit $\varepsilon \rightarrow 0$. Additionally, we improve many bounds for the case where the smoothing parameter is close to 1. This extension is important to show strong converse statements, as we will see in Section 8.1.

6.1.2 Outline

The remainder of the chapter is organized as follows. In Section 6.2, we explore the special case of classical registers without side information and introduce the proof techniques, based on relations between the smooth min-entropy and Rényi entropies, that will also lead to the fully quantum AEP.

¹Note also that the appearance of the reference system in the convergence rate is not entirely unexpected. This can be seen from the fact that (6.5) directly implies a bound on the max-entropy, i.e.

$$\frac{1}{n} H_{\max}^{\varepsilon}(A^n|R^n) \leq H(A|R) + \frac{\delta}{\sqrt{n}},$$

by duality of the entropies. Hence, it is no surprise that this duality also appears in the convergence rate.

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

In Section 6.3, we introduce quantum generalizations of the classical Rényi α -entropies. (See also Appendix B, where some of their properties are discussed.) Then, in Section 6.4, we show, analogous to the classical proof, that the smooth min-entropy can be bounded in terms of these entropies. Section 6.5 then bounds the difference between Rényi α -entropies and the von Neumann entropies for α close to 1. Finally, these results are combined in Section 6.6 to prove the fully quantum AEP.

6.2 Sketch of a Classical Proof

The relation between the traditional formulation of the AEP in terms of probabilities [CT91] and its entropic formulation was already explored in the preceding section. To give an intuitive idea of the techniques used in the following, we first prove a special case of the AEP for classical registers and take an additional limit $\varepsilon \rightarrow 0$. More precisely, we show that

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^{\varepsilon}(X^n) = H(X).$$

We employ the Rényi α -entropies [Rén61]

$$H_{\alpha}(X) := \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} P(x)^{\alpha}, \quad \alpha \in (0, 1) \cup (1, \infty), \quad (6.6)$$

for which H_{∞} ($\alpha \rightarrow \infty$), H_0 ($\alpha \rightarrow 0$) and the Shannon entropy ($\alpha \rightarrow 1$) are defined as limits. Furthermore, the entropies H_{α} are monotonically decreasing in α and, as shown in [RW04], the Rényi entropies with $\alpha > 1$ are close to the smooth min-entropy in the sense that

$$H_{\min}^{\varepsilon}(X) \geq H_{\alpha}(X) - \frac{1}{\alpha-1} \log \frac{1}{\varepsilon}, \quad \alpha > 1$$

while those with $\alpha < 1$ are close to the smooth max-entropy. Note that the error term $\frac{1}{\alpha-1} \log 1/\varepsilon$ in (6.7) diverges when we try to recover the Shannon entropy. However, in the case of an i.i.d. sequence we find

$$\frac{1}{n} H_{\min}^{\varepsilon}(X^n) \geq H_{\alpha}(X) - \frac{1}{n(\alpha-1)} \log \frac{1}{\varepsilon}, \quad (6.7)$$

where we have used $H_{\alpha}(X^n) = nH_{\alpha}(X)$. We proceed by bounding the entropy $\frac{1}{n} H_{\min}^{\varepsilon}(X^n)$ in the limit $n \rightarrow \infty, \varepsilon \rightarrow 0$ from above and below. To get the lower bound, we choose $\alpha = 1 + 1/\sqrt{n}$ and take the limit $n \rightarrow \infty$ in (6.7). An upper bound in this limit follows directly from the fact that $H_{\min}(X) \leq H(X)$ (cf. Proposition 4.3 or [Rén61]) and the continuity of the Shannon entropy.¹

A similar argument shows that the smooth max-entropy converges to the von Neumann entropy in the asymptotic limit.

¹See, e.g. Fannes [Fan73, AF04], where the relevant continuity property is shown for (conditional) von Neumann entropies.

6.3 Quantum Relative Entropies

We prove the asymptotic equipartition property for relative entropies, which are introduced here. Conditional entropies can be seen as special cases of relative entropies.

6.3.1 Quasi-Entropies

A very general class of classical relative entropies are the f -divergences, originally introduced by Csiszár [Csi72]. They have been generalized to the quantum setting by Petz [Pet84], who calls them quasi-entropies. (See also [OP93, Pet10, HMPB11] for an overview of recent results on these entropies.) The crucial observation is that some of the most interesting mathematical properties of the *von Neumann* and *Shannon* entropies are a direct consequence of the (operator) concavity of the function $h : t \mapsto -t \log t$ that defines the functional $H(\rho) = \text{tr}(h(\rho))$. Hence, the following generalization of this functional was investigated.

Definition 6.1 (Quasi-Entropy). *Let $A, B \in \mathcal{P}(\mathcal{H})$ and let $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ be continuous. Then, the f -quasi-entropy of A relative to B is*

$$S_f(A \| B) := \lim_{\xi \rightarrow 0} \left\langle \Gamma \left| \sqrt{B + \xi \mathbb{1}} f \left((B + \xi \mathbb{1})^{-1} \otimes A^T \right) \sqrt{B + \xi \mathbb{1}} \right| \Gamma \right\rangle,$$

where $|\Gamma\rangle = \sum_i |i\rangle \otimes |i\rangle$ and $\{|i\rangle\}$ is an orthonormal basis of \mathcal{H} , with respect to which the transpose is defined.

Unconditional quasi-entropies are recovered when we substitute $B = \mathbb{1}$. Let us consider a few examples of such functions. For this purpose, let $A, B \in \mathcal{P}(\mathcal{H})$ be two positive operators whose supports satisfy $\text{supp}\{A\} \subseteq \text{supp}\{B\}$, such that the limit in Def. 6.1 is always finite and B^{-1} can be interpreted as a *generalized inverse*. First, using $h : t \mapsto -t \log t$ extended to \mathbb{R}_0^+ using $h(0) = \lim_{t \rightarrow 0} h(t) = 0$, the relative von Neumann entropy is given as

$$S(A \| B) := S_h(A \| B) = \text{tr} (A(\log B - \log A)). \quad (6.8)$$

Note that we omit a minus sign here that is present in the conventional definition of the relative entropy. To derive this well-known expression from Def. 6.1, we used that $\langle \Gamma | X \otimes \mathbb{1} | \Gamma \rangle = \text{tr}(X)$ and $(X \otimes \mathbb{1}) | \Gamma \rangle = (\mathbb{1} \otimes X^T) | \Gamma \rangle$ for any operator $X \in \mathcal{L}(\mathcal{H})$.

6.3.2 Relative Rényi entropies

As another example, we recover generalizations of the Rényi entropies of order $\alpha \in (0, 1) \cup (1, \infty)$. This is achieved using $g_\alpha : t \mapsto t^\alpha$ and

$$S_\alpha(A \| B) := \frac{1}{1 - \alpha} \log S_{g_\alpha}(A \| B) = \frac{1}{1 - \alpha} \log \text{tr}(A^\alpha B^{1-\alpha}), \quad (6.9)$$

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

where, for $\alpha > 1$, we use the generalized inverse of σ . We may continuously extend the range of α to the limits $\alpha \rightarrow 0$ and $\alpha \rightarrow \infty$. Moreover, the von Neumann relative entropy emerges in the limit $\alpha \rightarrow 1^\pm$ if $\text{tr}(A) = 1$. Hence, we can continuously extend the range of valid parameters to $\alpha \in \mathbb{R}_0^+$ by setting $S_1 \equiv S$ in this case. Some properties of these entropies are discussed in Appendix B.

6.3.3 Relative Min- and Max-Entropy

In addition to this, we will also need a relative entropy version of the min- and max-entropies. The relative min-entropy was introduced by Datta [Dat09]¹ and is directly related to the conditional min-entropy defined in [Ren05]. The relative min- and max-entropies of $A \in \mathcal{P}(\mathcal{H})$ relative to $B \in \mathcal{P}(\mathcal{H})$ are given by

$$\begin{aligned} S_{\min}(A \| B) &:= \sup \{ \lambda : A \leq 2^{-\lambda} B \} \quad \text{and} \\ S_{\max}(A \| B) &:= \log \left\| \sqrt{A} \sqrt{B} \right\|_1^2. \end{aligned} \quad (6.10)$$

We also need a smoothed version of the relative min-entropy, which we define for any $\rho \in \mathcal{S}_\leq(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$.

$$S_{\min}^\varepsilon(\rho \| \sigma) := \max_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)} S_{\min}(\tilde{\rho} \| \sigma). \quad (6.11)$$

Clearly, the (smooth) min- and the max-entropy (as defined in Chapter 4) of A conditioned on B , $H_{\min}^\varepsilon(A|B)_\rho$ and $H_{\max}(A|B)_\rho$, can be recovered by the substitution $A = \rho_{AB}$ and $B = \mathbb{1}_A \otimes \sigma_B$, where σ_B is maximized over $\mathcal{S}_=(\mathcal{H}_B)$.

6.4 Lower Bounds on the Smooth Min-Entropy

The following lemma gives a first lower bound on the smooth relative entropy (6.11). (A similar, less general lemma can be found in [DR09].)

Lemma 6.1. *Let $\rho \in \mathcal{S}_\leq(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$ and $\lambda \geq S_{\min}(\rho \| \sigma)$. Then,*

$$S_{\min}^\varepsilon(\rho \| \sigma) \geq \lambda, \quad \text{where } \varepsilon = \sqrt{2 \text{tr}(\Delta) - \text{tr}(\Delta)^2} \quad \text{and } \Delta = \{\rho - 2^{-\lambda} \sigma\}_+.$$

Proof. We first choose $\tilde{\rho}$, bound $S_{\min}^\varepsilon(\tilde{\rho} \| \sigma)$, and then show that $\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)$. We use the abbreviated notation $\Lambda := 2^{-\lambda} \sigma$ and set

$$\tilde{\rho} := G \rho G^\dagger, \quad \text{where } G := \Lambda^{\frac{1}{2}} (\Lambda + \Delta)^{-\frac{1}{2}},$$

where we use the *generalized inverse*. From the definition of Δ , we have $\rho \leq \Lambda + \Delta$; hence, $\tilde{\rho} \leq \Lambda$ and $S_{\min}(\tilde{\rho} \| \sigma) \geq \lambda$.

¹There, it appears under the name $D_{\max} \equiv -S_{\min}$.

6.4 Lower Bounds on the Smooth Min-Entropy

Let $|\psi\rangle$ be a purification of ρ , then $(G \otimes \mathbb{1})|\psi\rangle$ is a purification of $\tilde{\rho}$ and, using Uhlmann's theorem (3.3), we find a bound on the (generalized) fidelity:

$$\begin{aligned} F(\tilde{\rho}, \rho) &\geq |\langle \psi | G | \psi \rangle| + \sqrt{(1 - \text{tr}(\rho))(1 - \text{tr}(\tilde{\rho}))} \\ &\geq \Re\{\text{tr}(G\rho)\} + 1 - \text{tr}(\rho) = 1 - \text{tr}((\mathbb{1} - \bar{G})\rho), \end{aligned}$$

where we introduced $\bar{G} := \frac{1}{2}(G + G^\dagger)$. This can be simplified further after we note that G is a contraction.¹ To see this, we multiply $\Lambda \leq \Lambda + \Delta$ with $(\Lambda + \Delta)^{-\frac{1}{2}}$ from left and right to get

$$G^\dagger G = (\Lambda + \Delta)^{-\frac{1}{2}} \Lambda (\Lambda + \Delta)^{-\frac{1}{2}} \leq \mathbb{1}. \quad (6.12)$$

Furthermore, $\bar{G} \leq \mathbb{1}$, since $\|\bar{G}\| \leq 1$ by the triangle inequality and $\|G\| = \|G^\dagger\| \leq 1$. Moreover,

$$\begin{aligned} \text{tr}((\mathbb{1} - \bar{G})\rho) &\leq \text{tr}(\Lambda + \Delta) - \text{tr}(\bar{G}(\Lambda + \Delta)) \\ &= \text{tr}(\Lambda + \Delta) - \text{tr}((\Lambda + \Delta)^{1/2} \Lambda^{1/2}) \leq \text{tr}(\Delta), \end{aligned}$$

where we used $\rho \leq \Lambda + \Delta$ and $\sqrt{\Lambda + \Delta} \geq \sqrt{\Lambda}$. The latter inequality follows from the operator monotonicity of the square root function (cf. Section 2.3.2). Finally, using the above bounds, the purified distance between $\tilde{\rho}$ and ρ is bounded by

$$P(\tilde{\rho}, \rho) = \sqrt{1 - F^2(\tilde{\rho}, \rho)} \leq \sqrt{1 - (1 - \text{tr}(\Delta))^2} = \sqrt{2 \text{tr}(\Delta) - \text{tr}(\Delta)^2}.$$

Hence, we verified that $\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)$, which concludes the proof. \square

In particular, this means that for a fixed $\varepsilon \in [0, 1)$ and $\text{supp}\{\rho\} \subseteq \text{supp}\{\sigma\}$, we can always find a finite λ s.t. Lemma 6.1 holds. To see this, note that $\varepsilon(\lambda) = \sqrt{2 \text{tr}(\Delta) - \text{tr}(\Delta)^2}$ is continuous in λ with $\varepsilon(S_{\min}(\rho \parallel \sigma)) = 0$ and $\lim_{\lambda \rightarrow \infty} \varepsilon(\lambda) = 1$.

Rényi Entropies and Smooth Min-Entropy

Our main tool for proving the fully quantum AEP is a family of inequalities that relate the smooth relative min-entropy to relative Rényi entropies for $\alpha \in (1, 2]$. This family contains the von Neumann relative entropy in the limit $\alpha \rightarrow 1$. This can be seen as a quantum generalization of the classical inequality in [RW04].

Proposition 6.2. *Let $\rho \in \mathcal{S}_\leq(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$, $0 < \varepsilon < 1$ and $\alpha \in (1, 2]$. Then,*

$$S_{\min}^\varepsilon(\rho \parallel \sigma) \geq S_\alpha(\rho \parallel \sigma) - \frac{g(\varepsilon)}{\alpha - 1}, \quad \text{where } g(\varepsilon) = \log \frac{1}{1 - \sqrt{1 - \varepsilon^2}}. \quad (6.13)$$

¹A contraction G is an operator with operator norm $\|G\| \leq 1$.

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

Proof. We consider two cases: (1) The α entropy diverges (to $-\infty$) and the inequality holds trivially. (2) We have $\text{supp } \{\rho\} \subseteq \text{supp } \{\sigma\}$. In this case, we can find an isometry $\mathcal{H}' \rightarrow \mathcal{H}$ that maps a σ' to σ and ρ' to ρ s.t. σ' has full support. The min- and α -entropies are invariant under this isometry due to Proposition 5.3 and Lemma B.1, thus, we henceforth assume that σ is invertible in this proof.

We use Lemma 6.1 to get a first bound on D_{\min}^ε ; in particular, we choose λ s.t. Lemma 6.1 holds for ε . Next, we introduce the operator $X := \rho - 2^{-\lambda}\sigma$ with eigenbasis $\{|e_i\rangle\}_{i \in S}$. The set $S^+ \subseteq S$ contains the indices i corresponding to positive eigenvalues of X . Hence, $P^+ := \sum_{i \in S^+} |e_i\rangle\langle e_i|$ is the projector on the positive eigenspace of X and $P^+ X P^+ = \Delta$ as defined in Lemma 6.1. Furthermore, let $r_i := \langle e_i | \rho | e_i \rangle \geq 0$ and $s_i := \langle e_i | \sigma | e_i \rangle > 0$. It follows that

$$\forall i \in S^+ : r_i - 2^{-\lambda} s_i \geq 0 \quad \text{and, thus,} \quad \frac{r_i}{s_i} 2^\lambda \geq 1.$$

For any $\alpha \in (1, 2]$, we bound $\text{tr}(\Delta) = 1 - \sqrt{1 - \varepsilon^2}$ as follows:

$$\begin{aligned} 1 - \sqrt{1 - \varepsilon^2} = \text{tr}(\Delta) &= \sum_{i \in S^+} r_i - 2^{-\lambda} s_i \leq \sum_{i \in S^+} r_i \\ &\leq \sum_{i \in S^+} r_i \left(\frac{r_i}{s_i} 2^\lambda \right)^{\alpha-1} \leq 2^{\lambda(\alpha-1)} \sum_{i \in S} r_i^\alpha s_i^{1-\alpha}. \end{aligned}$$

Hence, taking the logarithm and dividing by $\alpha - 1 > 0$, we get

$$\lambda \geq \frac{1}{1-\alpha} \log \sum_{i \in S} r_i^\alpha s_i^{1-\alpha} - \frac{1}{\alpha-1} \log \frac{1}{1 - \sqrt{1 - \varepsilon^2}}. \quad (6.14)$$

Next, we use the monotonicity of the Rényi entropies (cf. Lemma B.4). We use the measurement TP-CPM $\mathcal{M} : X \mapsto \sum_{i \in S} |e_i\rangle\langle e_i| X |e_i\rangle\langle e_i|$ to obtain

$$S_\alpha(\rho \| \sigma) \leq S_\alpha(\mathcal{M}(\rho) \| \mathcal{M}(\sigma)) = \frac{1}{1-\alpha} \log \sum_{i \in S} r_i^\alpha s_i^{1-\alpha}.$$

We conclude the proof by substituting this into (6.14) and applying the upper bound on λ in Lemma 6.1. \square

We also note here that the term $1 - \sqrt{1 - \varepsilon^2}$ in $g(\varepsilon)$ can be bounded by simpler expressions (cf. Fig. 6.2). We find $1 - \sqrt{1 - \varepsilon^2} \geq \frac{\varepsilon^2}{2}$ using a second order Taylor expansion of the expression around $\varepsilon = 0$ and the fact that the third derivative is non-negative. This is a very good approximation for small ε . Hence, Eq. (6.13) can be simplified to

$$S_{\min}^\varepsilon(\rho \| \sigma) \geq S_\alpha(\rho \| \sigma) - \frac{1}{\alpha-1} \log \frac{2}{\varepsilon^2}. \quad (6.15)$$

This form of the inequality been reported previously [TCR09].

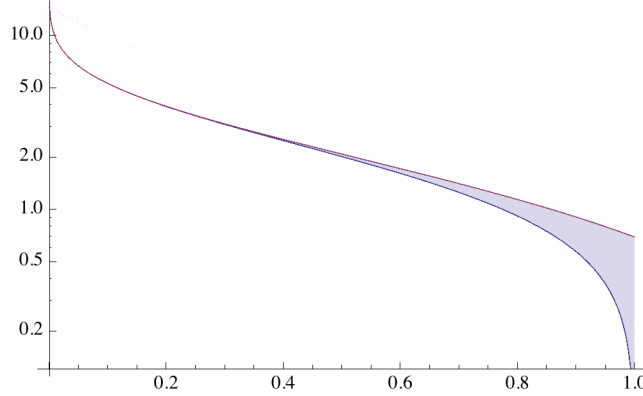


Figure 6.2: *Bound on $g(\varepsilon)$.* Plot of the function $g(\varepsilon) = -\log(1 - \sqrt{1 - \varepsilon^2})$ and its upper bound, $\log \frac{2}{\varepsilon^2}$. The horizontal axis denotes ε .

Conditional Entropies

This lemma is of particular interest when applied to the smooth conditional min-entropy. In this case, let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and σ be of the form $\mathbb{1}_A \otimes \rho_B$. Then, for any $\alpha \in (1, 2]$, we have

$$\begin{aligned} H_{\min}^{\varepsilon}(A|B)_{\rho} &\geq S_{\min}^{\varepsilon}(\rho_{AB}, \mathbb{1}_A \otimes \rho_B) \\ &\geq S_{\alpha}(\rho_{AB}, \mathbb{1}_A \otimes \rho_B) - \frac{g(\varepsilon)}{\alpha - 1} = H_{\alpha}(A|B)_{\rho} - \frac{g(\varepsilon)}{\alpha - 1}. \end{aligned} \quad (6.16)$$

The duality relation for the smooth min- and max-entropies (cf. Theorem 5.4) and the Rényi α -Entropies (cf. Lemma B.5) imply a corresponding dual relation for the max-entropy. For any $\alpha \in [0, 1)$, we have

$$H_{\max}^{\varepsilon}(A|B)_{\rho} \leq H_{\alpha}(A|B)_{\rho} - \frac{g(\varepsilon)}{1 - \alpha}. \quad (6.17)$$

6.5 From Rényi to von Neumann Entropies

We will use Proposition 6.2 to get a lower bound on the smooth min-entropy in terms of α -entropies, hence, it remains to find a lower bound on the α -entropies in terms of the von Neumann entropy. In turn, the bound on the convergence will depend on the smoothing parameter ε and a contribution $\Upsilon(\rho \parallel \sigma)$ that describes how fast the α -entropies converge to the von Neumann entropy.

Definition 6.2. Let $\rho, \sigma \in \mathcal{P}(\mathcal{H})$, then we define the α -entropy convergence parameter,

$$\Upsilon(\rho \parallel \sigma) := 2^{-\frac{1}{2}S_{3/2}(\rho \parallel \sigma)} + 2^{\frac{1}{2}S_{1/2}(\rho \parallel \sigma)} + 1.$$

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

We can now state a bound on the α -entropies for α close to 1.

Lemma 6.3. *Let $\rho \in \mathcal{S}_=(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$ and $1 < \alpha < 1 + \frac{\log 3}{4 \log v}$, where $v = \Upsilon(\rho \parallel \sigma)$. Then,*

$$S_\alpha(\rho \parallel \sigma) > S(\rho \parallel \sigma) - 4(\alpha - 1)(\log v)^2.$$

Proof. We assume that σ is invertible in this proof. The general result then follows by the arguments outlined at the beginning of the proof of Proposition 6.2.

Let $\{|i\rangle\}$ be an orthonormal basis of \mathcal{H} . The state $|\gamma\rangle := \sum_i |i\rangle \otimes |i\rangle$ is the (unnormalized) fully entangled state on $\mathcal{H} \otimes \mathcal{H}$. We introduce a purification $|\varphi\rangle := \sqrt{\rho} |\gamma\rangle$ of ρ . To simplify notation, we use $\beta := \alpha - 1$ as well as $X := \rho \otimes (\sigma^{-1})^T$.

Let us first approximate S_α for small $\beta > 0$.

$$S_\alpha(\rho \parallel \sigma) = -\frac{1}{\beta} \log \langle \varphi | X^\beta | \varphi \rangle \geq \frac{1}{\beta \ln 2} (1 - \langle \varphi | X^\beta | \varphi \rangle),$$

where we used $\ln x \leq x - 1$ for all $x > 0$. We now expand the exponential t^β for each eigenvalue $t > 0$ of X as follows: $t^\beta = 1 + \beta \ln t + r_\beta(t)$, where $r_\beta(t) := t^\beta - \beta \ln t - 1$. This leads to

$$\begin{aligned} S_\alpha(\rho \parallel \sigma) &\geq \frac{1}{\beta \ln 2} (1 - \text{tr}(\rho) - \beta \langle \varphi | \ln X | \varphi \rangle - \langle \varphi | r_\beta(X) | \varphi \rangle) \\ &\geq S(\rho \parallel \sigma) - \frac{1}{\beta \ln 2} \langle \varphi | r_\beta(X) | \varphi \rangle. \end{aligned} \quad (6.18)$$

To simplify this further, we note that

$$r_\beta(t) \leq 2(\cosh(\beta \ln t) - 1) =: s_\beta(t).$$

It is easy to verify that s_β is monotonically increasing for $t \geq 1$ and concave in t for $\beta \leq 1/2$ and $t \in [3, \infty)$. Furthermore, we have $s_\beta(t) = s_\beta(\frac{1}{t})$ and $s_\beta(t^2) = s_{2\beta}(t)$. We use this to bound¹

$$s_\beta(t) \leq s_\beta\left(t + \frac{1}{t} + 2\right) = s_{2\beta}\left(\sqrt{t} + \frac{1}{\sqrt{t}}\right) \leq s_{2\beta}\left(\sqrt{t} + \frac{1}{\sqrt{t}} + 1\right). \quad (6.19)$$

Next, we apply (6.19) to the matrix element in (6.18) and use the fact that the operator $\sqrt{X} + 1/\sqrt{X} + \mathbb{1}$ has its eigenvalues in $[3, \infty)$ and $2\beta < \frac{\log 3}{2 \log v} \leq \frac{1}{2}$ together with Lemma 2.7 to get

$$\langle \varphi | s_\beta(X) | \varphi \rangle \leq \langle \varphi | s_{2\beta}\left(\sqrt{X} + \frac{1}{\sqrt{X}} + \mathbb{1}\right) | \varphi \rangle \leq s_{2\beta}(v), \quad (6.20)$$

¹Adaptions of this step lead to different bounds. Here, we are interested in a bound that can be expressed in terms of $S_{1/2}$ and $S_{3/2}$.

6.5 From Rényi to von Neumann Entropies

where we substituted $v = \langle \varphi | \sqrt{X} + 1/\sqrt{X} + \mathbb{1} | \varphi \rangle$.

Taylor's theorem and an expansion around $\beta = 0$ gives an upper bound on $s_\beta(t)$: $s_\beta(t) \leq \beta^2 (\ln t)^2 \cosh(\beta \ln t)$. Hence,

$$\frac{1}{\beta \ln 2} s_{2\beta}(v) \leq 4\beta (\log v)^2 \ln 2 \cosh(2\beta \ln v) < 4\beta (\log v)^2, \quad (6.21)$$

where we simplified the expression (for convenience of exposition) using $\ln 2 \cosh(\ln 3/2) < 1$. The lemma now follows after we substitute (6.21) and (6.20) into (6.18). \square

This Lemma can be extended to include sub-normalized states, $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$. Let $\hat{\rho} = \rho / \text{tr}(\rho)$ and $\hat{v} = \Upsilon(\hat{\rho} \parallel \sigma)$, then

$$\begin{aligned} S_\alpha(\rho \parallel \sigma) &= S_\alpha(\hat{\rho} \parallel \sigma) + \frac{\alpha}{\alpha - 1} \log \frac{1}{\text{tr}(\rho)} \\ &> S(\hat{\rho} \parallel \sigma) + \frac{\alpha}{\alpha - 1} \log \frac{1}{\text{tr}(\rho)} - 4(\alpha - 1)(\log \hat{v})^2 \\ &= \frac{1}{\text{tr}(\rho)} S(\rho \parallel \sigma) + \frac{1}{\alpha - 1} \log \frac{1}{\text{tr}(\rho)} - 4(\alpha - 1)(\log \hat{v})^2. \end{aligned}$$

Now we combine Proposition 6.2 and 6.3 to get the desired bound for i.i.d. operators. (Note that we restrict ρ to normalized states because, if $\text{tr}(\rho) < 1$ then the trace of $\rho^{\otimes n}$ drops exponentially to zero. An extension of this to sub-normalized states, while possible, thus seems uninteresting.)

Theorem 6.4. *Let $\rho \in \mathcal{S}_{=}(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$, $0 < \varepsilon < 1$ and $v = \Upsilon(\rho \parallel \sigma)$. Then, for any $n \geq \frac{8}{5}g(\varepsilon)$, the i.i.d. operators $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ satisfy*

$$\frac{1}{n} S_{\min}^\varepsilon(\rho^{\otimes n} \parallel \sigma^{\otimes n}) \geq S(\rho \parallel \sigma) - \frac{\delta(\varepsilon, v)}{\sqrt{n}} \quad \text{where} \quad \delta(\varepsilon, v) = 4 \log v \sqrt{g(\varepsilon)}$$

and $g(\varepsilon) = -\log(1 - \sqrt{1 - \varepsilon^2})$.

Proof. We use Proposition 6.2, the additivity property of the Rényi entropy (B.3) and Lemma 6.3 to get a bound on the smooth min-entropy. Let $\alpha := 1 + \frac{1}{2\mu\sqrt{n}}$ for a parameter μ (to be optimized over), then

$$\begin{aligned} \frac{1}{n} S_{\min}^\varepsilon(\rho^{\otimes n} \parallel \sigma^{\otimes n}) &\geq \frac{1}{n} S_\alpha(\rho^{\otimes n} \parallel \sigma^{\otimes n}) - \frac{1}{n(\alpha - 1)} g(\varepsilon) \\ &= S_\alpha(\rho \parallel \sigma) - \frac{2\mu}{\sqrt{n}} g(\varepsilon) \\ &\geq S(\rho \parallel \sigma) - \frac{2}{\sqrt{n}} \left(\mu g(\varepsilon) + \frac{1}{\mu} (\log v)^2 \right). \end{aligned} \quad (6.22)$$

Clearly, we want to choose μ such that it minimizes the expression $\mu g(\varepsilon) + \mu^{-1} (\log v)^2$. However, the requirement $\alpha < 1 + \frac{\log 3}{4 \log v}$ in Lemma 6.3 restricts

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

the choice of μ for any fixed n , hence, the error term is in general also a function of n . Nonetheless, for large enough n the optimum, μ_* , can be reached¹ and we get

$$\mu_* = \sqrt{\frac{(\log v)^2}{g(\varepsilon)}} \quad \text{for} \quad n \geq \frac{8}{5} \frac{(\log v)^2}{\mu_*^2} = \frac{8}{5} g(\varepsilon). \quad (6.23)$$

Substitution of this expression into (6.22) concludes the proof. \square

6.6 The Asymptotic Equipartition Property

6.6.1 Direct Part

In this section, we are mostly interested in the application of Theorem 6.4 to conditional min- and max-entropies. Here, for any state $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$, we chose $\sigma_{AB} = \mathbb{1}_A \otimes \rho_B$ and apply Theorem 6.4. This implies that

$$\begin{aligned} \frac{1}{n} H_{\min}^\varepsilon(A^n|B^n)_\rho &\geq \frac{1}{n} S_{\min}^\varepsilon(\rho_{AB}^{\otimes n}, \sigma_{AB}^{\otimes n}) \\ &\geq S(\rho_{AB}, \sigma_{AB}) - \frac{\delta(\varepsilon, v)}{\sqrt{n}} = H(A|B)_\rho - \frac{\delta(\varepsilon, v)}{\sqrt{n}}. \end{aligned}$$

This (and the dual of this relation) leads to the following corollary.

Corollary 6.5. *Let $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$ and $0 < \varepsilon < 1$. Then, the smooth entropies of the i.i.d. product state $\rho_{A^n B^n} = \rho_{AB}^{\otimes n}$ satisfy*

$$\frac{1}{n} H_{\min}^\varepsilon(A^n|B^n)_\rho \geq H(A|B)_\rho - \frac{\delta(\varepsilon, v)}{\sqrt{n}} \quad \text{and} \quad (6.24)$$

$$\frac{1}{n} H_{\max}^\varepsilon(A^n|B^n)_\rho \leq H(A|B)_\rho + \frac{\delta(\varepsilon, v)}{\sqrt{n}}, \quad (6.25)$$

where $\delta(\varepsilon, v)$ is defined in Theorem 6.4 and $v = \Upsilon(\rho_{AB} \|\mathbb{1}_A \otimes \rho_B)$.

We may now use (B.2) and its dual relation, $H_{3/2}(A|B)_\rho \geq H_{\min}(A|B)_\rho$ (cf. Theorem 5.4 and Lemma B.5), to bound

$$\Upsilon(\rho_{AB} \|\mathbb{1}_A \otimes \rho_B) \leq \sqrt{2^{-H_{\min}(A|B)_\rho}} + \sqrt{2^{H_{\max}(A|B)_\rho}} + 1.$$

This is Result 5 of Section 6.1.

The following is a trivial corollary from Theorem 6.4 and the above arguments, in particular (6.24) and its dual relation (6.25).

Corollary 6.6 (AEP, direct). *Let $\rho \in \mathcal{S}_=(\mathcal{H}_{AB})$ and $0 < \varepsilon < 1$. Then,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^\varepsilon(A^n|B^n)_\rho \geq H(A|B)_\rho \geq \lim_{n \rightarrow \infty} \frac{1}{n} H_{\max}^\varepsilon(A^n|B^n)_\rho.$$

¹To verify this, evaluate an upper bound to $\alpha = 1 + (2\mu_*\sqrt{n})^{-1}$ using the expression for n in (6.23) and note that $\sqrt{5/2} < \log 3$.

6.6.2 Converse Part

To prove asymptotic convergence, we will also need converse bounds. For $\varepsilon = 0$, the converse bounds are given by Proposition 4.3, i.e. $H_{\min}(A|B) \leq H(A|B) \leq H_{\max}(A|B)$. For $\varepsilon > 0$, similar bounds can be derived from the continuity of the conditional entropy in the state [AF04]. However, such bounds do not allow a statement of the form of Corollary 6.6 as the deviation from the von Neumann entropy scales as $nf(\varepsilon)$, where $f(\varepsilon) \rightarrow 0$ only for $\varepsilon \rightarrow 0$. (See, for example, [TCR09] for such a weak converse bound.) This is not sufficient for some applications of the asymptotic equipartition property.

Here, we prove a tighter bound, which relies on the bound between smooth max-entropy and smooth min-entropy established in Proposition 5.5. Applying this Proposition in conjunction with Eqs. (6.24) and (6.25) establishes the converse AEP bounds. Let $0 < \varepsilon < 1$. Then, using any smoothing parameter $0 < \varepsilon' < 1 - \varepsilon$, we bound

$$\begin{aligned} \frac{1}{n} H_{\min}^{\varepsilon}(A|B)_{\rho} &\leq \frac{1}{n} H_{\max}^{\varepsilon'}(A|B)_{\rho} + \frac{1}{n} \log \frac{1}{1 - (\varepsilon + \varepsilon')^2} \\ &\leq H(A|B)_{\rho} + \frac{1}{n} \log \frac{1}{1 - (\varepsilon + \varepsilon')^2} + \frac{\delta(\varepsilon', \eta)}{\sqrt{n}}. \end{aligned} \quad (6.26)$$

The corresponding statement for the smooth max-entropy follows by the dual argument. We thus find

Corollary 6.7 (AEP, converse). *Let $\rho \in \mathcal{S}_{=}(H_{AB})$ and $0 < \varepsilon < 1$. Then,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^{\varepsilon}(A^n|B^n)_{\rho} \leq H(A|B)_{\rho} \leq \lim_{n \rightarrow \infty} \frac{1}{n} H_{\max}^{\varepsilon}(A^n|B^n)_{\rho}.$$

These converse bounds are particularly important to bound the smooth entropies for large smoothing parameters. In this form, the AEP implies strong converse statements for many information theoretic tasks that can be characterized by smooth entropies in the one-shot setting (see, for example, Chapter 8).

6. THE QUANTUM ASYMPTOTIC EQUIPARTITION PROPERTY

Chapter 7

Uncertainty Relations for Smooth Entropies

Entropic uncertainty relations (UCRs) provide lower bounds on the uncertainty of the outcomes of two incompatible measurements given side information in terms of conditional entropies. In this chapter, which is based on [TR11] and [HT11], we prove several uncertainty relations using *smooth min- and max-entropies* as well as the *von Neumann entropy* as measures of uncertainty.

7.1 Introduction and Related Work

Uncertainty relations have inspired physicists since the early days of quantum mechanics, when Heisenberg [Hei27] formulated his famous uncertainty principle. In its Robertson [Rob29] form, it states that the product of the standard deviations of the outcomes of two *incompatible measurements* on a pure state $|\psi\rangle$ is lower bounded in terms of the commutator of the *observables* (\hat{X} and \hat{Y}) of these measurements,

$$\sigma_X \cdot \sigma_Y \geq \frac{1}{2} \left| \langle \psi | [\hat{X}, \hat{Y}] | \psi \rangle \right|. \quad (7.1)$$

Here, the variance of the measurement outcomes is given as $\sigma_Z^2 = \langle \psi | \hat{Z}^2 | \psi \rangle - \langle \psi | \hat{Z} | \psi \rangle^2$ where Z is either X or Y . The commutator, $[\hat{X}, \hat{Y}] = \hat{X}\hat{Y} - \hat{Y}\hat{X}$, quantifies the incompatibility of the two observables.

Note that the uncertainty relation becomes trivial if the measured state is an eigenstate of either X or Y . More generally, the lower bound in (7.1) depends on the state $|\psi\rangle$ before measurement, which is often undesirable. For example, the state might be unknown or, in a cryptographic setting, prepared by an adversary. Furthermore, the standard deviation is not our preferred measure of uncertainty as it conflates two concepts: the value associated with different measurement outcomes and the uncertainty in

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

the probability distribution of the outcomes. The latter is the uncertainty we are interested in and is often quantified by entropies.

Uncertainty relations in terms of entropies were first proposed by Hirschman [Hir57], and Deutsch [Deu83] for the case of a finite output alphabet.¹ In the form proposed by Maassen and Uffink [MU88], it states that

$$H(X) + H(Y) \geq \log \frac{1}{c}, \quad \text{where} \quad c = \max_{x,y} |\langle x|y \rangle|^2. \quad (7.2)$$

for any state ρ before measurement. Two *registers*, X and Y , store the respective outcome of two different projective measurements, \mathcal{X} and \mathcal{Y} . The *overlap*, c , is determined by these measurements and independent of the state before measurement. More specifically, the maximization in the definition of the overlap is taken over all eigenvectors $|x\rangle$ of \hat{X} and $|y\rangle$ of \hat{Y} , where \hat{X} and \hat{Y} are observables corresponding to the projective measurements \mathcal{X} and \mathcal{Y} .

More generally, Krishna and Parthasarathy [KP02] considered *POVMs*, given by sets $\mathcal{X} = \{M_x\}$ and $\mathcal{Y} = \{N_y\}$ of positive semi-definite operators. For such measurements, the states of the registers containing the measurement outcomes are given by

$$\rho_X = \sum_x \text{tr} \left(\sqrt{M_x} \rho \sqrt{M_x} \right) |x\rangle\langle x| \quad \text{and} \quad \rho_Y = \sum_y \text{tr} \left(\sqrt{N_y} \rho \sqrt{N_y} \right) |y\rangle\langle y|.$$

The uncertainty relation (7.2) now holds for these states and the overlap

$$c := \max_{x,y} \left\| \sqrt{M_x} \sqrt{N_y} \right\|_{\infty}^2. \quad (7.3)$$

Note that the overlap reduces to (7.2) if the measurements are projective.

The uncertainty relation can be extended to one with classical side information (see also [Hal95, CBKG02]). Let the state between the system to be measured, A , and an observer, O , be of *CQ* form $\rho_{AO} = \sum_o p_o |o\rangle\langle o| \otimes \rho_A^o$. This can be seen as the observer preparing the state ρ_A^o with probability p_o . What is the entropy of the observer O about the measurement outcomes X and Y ? It is easy to see that

$$H(X|O) + H(Y|O) = \sum_o p_o \left(H(X)_{\rho^o} + H(Y)_{\rho^o} \right) \geq \log \frac{1}{c}.$$

Thus, the uncertainty relation (7.2) still holds. Furthermore, we may model the random basis choice by a uniformly distributed bit $\Theta \in \{\mathcal{X}, \mathcal{Y}\}$ that is independent of the state ρ_{AO} . Hence, we consider a joint state $\rho_{AO\Theta} = \rho_{AO} \otimes \pi_{\Theta}$ and a measurement on the A system that depends on the classical

¹See also [WW10], which offers a comprehensive review of uncertainty relations.

register Θ .¹ The measurement outcome is denoted Z , which replaces X and Y . Thus,

$$H(Z|O\Theta) = \frac{1}{2}H(X|O) + \frac{1}{2}H(Y|O) \geq \frac{1}{2} \log \frac{1}{c}. \quad (7.4)$$

This uncertainty relation with classical side information is also visualized in Figure 7.1.

A statement of the uncertainty relation in this form naturally leads to the question of what happens if the observer is allowed to store quantum information about the system, i.e. if the state ρ_{AO} is entangled. For the case of (7.4), only the trivial bound $H(Z|O\Theta)_\rho \geq 0$ holds. To see this, consider a fully entangled state between A and O . In this case, the observer can predict the value of Z perfectly using an appropriate measurement (depending on Θ) on his system. This is explained in Figure 7.2.

However, due to the monogamy of entanglement², it is unclear what happens if we introduce a second observer. Note, for example, that the entropy $H(A|O)$ before measurement can be negative for entangled states while the sum $H(A|O_1) + H(A|O_2)$ is always non-negative.

In fact, Renes and Boileau [RB09] (see also [CW05]) conjectured the following uncertainty relation, which was later shown by Berta et al. [BCC⁺10]. For a tripartite quantum state shared between the system A which is measured in the basis determined by Θ and two observers, O_1 and O_2 , it holds that

$$H(Z|O_1\Theta) + H(Z|O_2\Theta) \geq \log \frac{1}{c}, \quad (7.5)$$

with c defined as in (7.3) but only for rank-1 projective measurements. This proof was later extended to POVMs and simplified by Coles et al. [CYGG11, CYZ11]. Concurrently, our results also imply a simplified proof of the uncertainty relation for von Neumann entropies and POVMs.

7.1.1 Main Contributions

We have already seen in Chapter 5 that the smooth entropies, for any tripartite state ρ_{ABC} and $\varepsilon \geq 0$, satisfy the duality relation

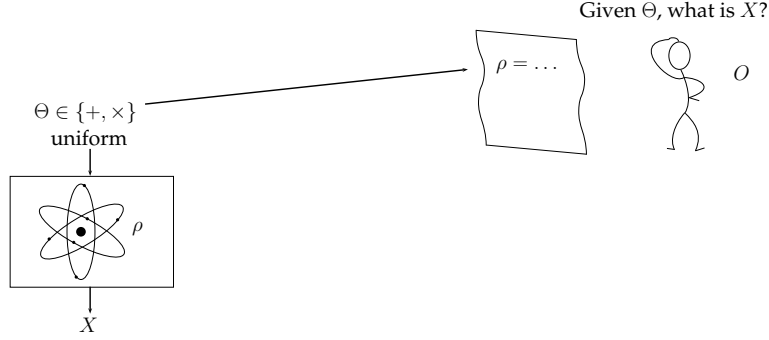
$$H_{\min}^\varepsilon(A|B)_\rho + H_{\max}^\varepsilon(A|C)_\rho \geq 0.$$

Our result now shows that this lower bound increases (as is the case for von Neumann entropies) when we apply incompatible measurements on the A system. This extends the entropic uncertainty relation of (7.5) to smooth entropies.

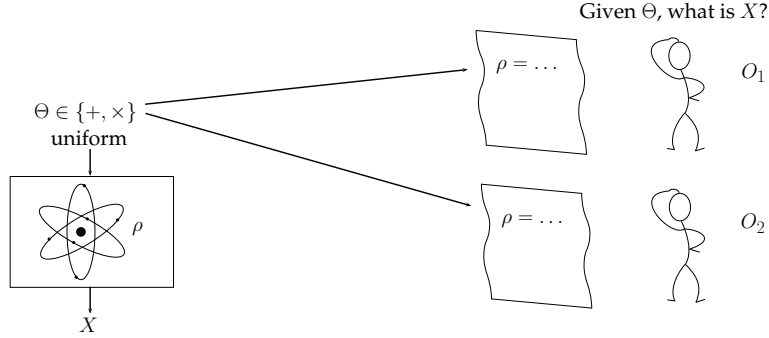
¹This can be seen a measurement of the joint system consisting of A and Θ .

²Or, depending on the reader's preference, due to no-cloning.

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

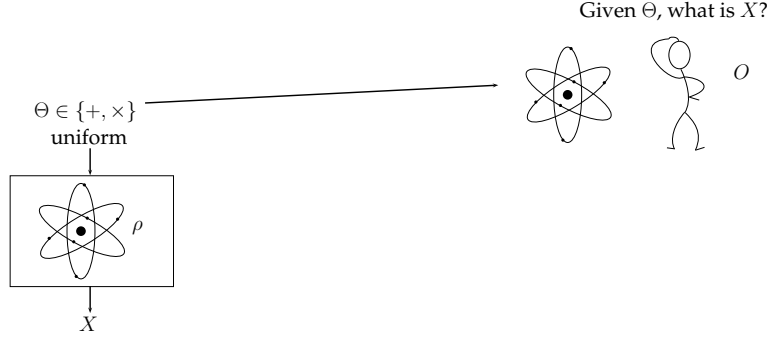


(a) A quantum system in the state ρ is measured in one out of two bases, determined by a uniformly random bit, Θ . Then the basis choice Θ is sent to an observer, O , who is asked to determine the measurement outcome, X . The observer is allowed to use classical information about the state ρ before measurement, including a full characterization of the density matrix. One may alternatively think of the observer preparing the state ρ before the game. The uncertainty relation (7.4) ensures that the entropy the observer has about the measurement outcome satisfies the lower bound $H(X|O\Theta) \geq \frac{1}{2} \log \frac{1}{c}$. Hence, if the overlap is nontrivial, the observer cannot predict X with certainty.

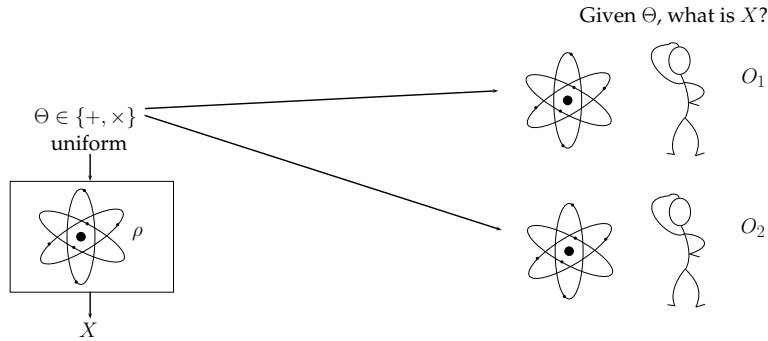


(b) This game can be trivially extended to two observers, O_1 and O_2 . In this case, the sum of their entropies satisfies $H(X|O_1\Theta) + H(X|O_2\Theta) \geq \log \frac{1}{c}$. This scenario is mainly interesting in the quantum case.

Figure 7.1: *Uncertainty Relations with Classical Side Information.*



(a) A quantum system, A , is measured in one out of two bases, determined by a uniformly random bit, Θ . Then the basis choice Θ is sent to an observer, O , who is asked to determine the measurement outcome, X . The observer holds a quantum system which might be entangled with the measured system. One may alternatively think of the observer preparing the bipartite state ρ_{AO} , of which the A part is measured and the O part is stored in a quantum memory. For this setup, no uncertainty relation of the type (7.4) exists. As a counter-example, consider the case where ρ_{AO} is fully entangled. In this case, the observer can always choose a measurement (depending on Θ) whose outcome is perfectly correlated with X . For example, if the state is a singlet, measuring both the A and O parts in the same, arbitrary basis will lead to perfectly anti-correlated binary variables.



(b) However, if we instead consider two observers, O_1 and O_2 , an uncertainty relation is possible. Note, in particular, that the counterexample provided for the case of one observer fails due to the monogamy of entanglement. As in the classical case, the sum of the entropies of the two observers satisfies $H(X|O_1\Theta) + H(X|O_2\Theta) \geq \log \frac{1}{c}$. However, in the quantum case it is possible that one of these entropies is zero, which implies that the other entropy is large. This trade-off between the two entropies does not exist in the classical case and can be seen as an effect of no-cloning—it is generally not possible that the two observers hold copies of each other's quantum information.

Figure 7.2: *Uncertainty Relations with Quantum Side Information.*

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

Result 7. For $\varepsilon \geq 0$, any tripartite state ρ_{ABC} as well as two POVMs $\{M_x\}$ and $\{N_y\}$ on A , it holds that

$$H_{\min}^{\varepsilon}(X|B) + H_{\max}^{\varepsilon}(Y|C) \geq \log \frac{1}{c}, \quad \text{where } c = \max_{x,y} \left\| \sqrt{M_x} \sqrt{N_y} \right\|_{\infty}^2.$$

This uncertainty relation has been used to prove security in quantum cryptography. In [TLGR11], it was used to derive tighter key rates for finite block lengths in quantum key distribution. In [LSS11], Le *et. al.* prove security of a reference frame independent quantum key distribution protocol using the above uncertainty relation.¹ Furthermore, the result has been shown in the framework of general von Neumann algebras in [BFS11], indicating that the UCR for min- and max-entropies is a fundamental property of quantum physics and not a relict of the density operator formalism.

Furthermore, we explore a tighter version of the uncertainty relation with a lower bound in terms of an *effective overlap*. The effective overlap—in contrast to the overlap—is a function of the marginal state prior to measurement as well as the two measurements. A preliminary version of this result appeared in [HT11], where the relation is shown for von Neumann entropies. We extend these results here and show that a generalized uncertainty relation also holds for smooth min- and max-entropies, enabling its substitution into existing cryptographic security proofs [TLGR11, LPT⁺11].

Result 8. For $\varepsilon > 0$, any tripartite state ρ_{ABC} as well as two POVMs $\mathcal{X} = \{M_x\}$ and $\mathcal{Y} = \{N_y\}$ on A , it holds that

$$H_{\min}^{3\varepsilon}(X|B) + H_{\max}^{\varepsilon}(Y|C) \geq \log \frac{1}{c^*} - \log \frac{2}{\varepsilon^2},$$

where $c^* = \sum_k \text{tr}(P^k \rho_A) c_k$. Here, $\{P^k\}$ is any projective measurement that commutes with both \mathcal{X} and \mathcal{Y} and c_k is the overlap of \mathcal{X} and \mathcal{Y} on the subspace P^k .

7.1.2 Outline

The remainder of this chapter is structured as follows. In Section 7.2, we introduce the notion of *overlap* and show the uncertainty relation for the min- and max-entropy. The proof is very instructive and will also guide the proof of the generalized uncertainty relation, which is given in Section 7.3. There, we will also formally define the effective overlap. In Section 7.4, we discuss a variety of corollaries of the generalized uncertainty relation. We consider an application of the uncertainty relation to quantum key distribution. Finally, in Section 7.5, we consider a bipartite uncertainty

¹Applications of the uncertainty relation to quantum cryptography will also be the topic of Section 8.3 of the next chapter.

relation which might have applications in cryptographic settings where only two parties are involved.

7.2 Traditional Formulation

7.2.1 Overlap

Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ be an arbitrary, tripartite quantum state. We want to bound the entropy about the result of a measurement on the A subsystem given side information stored in either B or C . Without loss of generality, such a measurement on the A system can be described by a POVM.¹ Here we consider two POVMs, $\mathcal{X} = \{M_x\}$ and $\mathcal{Y} = \{N_y\}$, on \mathcal{H}_A . They measure the state on the A subsystem and store the measurement outcome in classical registers X and Y , respectively. These two registers can be seen as classical random variables that are correlated with quantum side information on the B and C systems. Note that, in our analysis here, we are not concerned with the state of the A system after the measurement.

The relevant post-measurement states are thus given as

$$\rho_{XBC} = \sum_x |x\rangle\langle x| \otimes \text{tr}_A(\sqrt{M_x}\rho_{ABC}\sqrt{M_x}) \quad \text{and} \quad (7.6)$$

$$\rho_{YBC} = \sum_y |y\rangle\langle y| \otimes \text{tr}_A(\sqrt{N_y}\rho_{ABC}\sqrt{N_y}). \quad (7.7)$$

We also define the *overlap* of these two measurements.

Definition 7.1 (Overlap). *Let $\mathcal{X} = \{M_x\}$ and $\mathcal{Y} = \{N_y\}$ be two POVMs. Then, we define the overlap of \mathcal{X} and \mathcal{Y} as*

$$c(\mathcal{X}, \mathcal{Y}) := \max_{x,y} \|\sqrt{M_x}\sqrt{N_y}\|_{\infty}^2.$$

This is in accordance with [KP02], where an uncertainty relation (without side information) was first shown for von Neumann entropies. If the two measurements are projective, the expression for the overlap reduces to

$$c = \max_{x,y} |\langle x|y \rangle|^2,$$

where the maximization is over all eigenvectors $|x\rangle$ of \mathcal{X} and $|y\rangle$ of \mathcal{Y} . The name *overlap* is clearly motivated by this expression.

¹This is true since we do not use the post-measurement state on the A system and, thus, the freedom to choose a phase in the decomposition of the POVM elements $M = E^\dagger E$ is meaningless.

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

7.2.2 Uncertainty Relation for Min- and Max-Entropies

The following inequality (cf. [TR11]) bounds the sum of the min- and max-entropies of the post-measurement states in terms of the overlap. In the next sections, we will formulate a generalization of this statement to smooth entropies and mixed states. Here, we simply consider the special case (7.8), as its proof highlights the basic concepts and techniques used for the proof of the subsequent generalized uncertainty relation.

Theorem 7.1 (UCR). *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ be pure and \mathcal{X} as well as \mathcal{Y} two POVMs on \mathcal{H}_A . Then, the post measurement states (7.6) and (7.7) satisfy*

$$H_{\min}(X|B)_{\rho} + H_{\max}(Y|C)_{\rho} \geq \log \frac{1}{c(\mathcal{X}, \mathcal{Y})}. \quad (7.8)$$

Proof. It will be helpful to describe the two measurements in the Stinespring dilation picture (cf. Lemma 2.4) as isometries followed by a partial trace. Let U be the isometry from A to A, X and X' given by $U := \sum_x \sqrt{M_x} \otimes |x\rangle \otimes |x\rangle$. The isometry stores two copies of the measurement outcome in the registers X and X' and the measured state in A . Here, $\{|x\rangle\}$ is an orthonormal basis of $\mathcal{H}_X \cong \mathcal{H}_{X'}$. Analogously, $V := \sum_y \sqrt{N_y} \otimes |y\rangle \otimes |y\rangle$. Furthermore, we introduce the states $\rho_{AXX'BC} = U\rho_{ABC}U^{\dagger}$ and $\rho_{AYY'BC} = V\rho_{ABC}V^{\dagger}$, of which the post-measurement states appearing in Eq. (7.8), ρ_{XB} and ρ_{YC} , are marginals.

The duality relation (cf. Lemma 4.1) applied to $\tilde{\rho}_{AYY'BC}$ gives

$$H_{\max}(Y|C)_{\rho} + H_{\min}(Y|AY'B)_{\rho} = 0. \quad (7.9)$$

Comparing (7.9) with the statement of the theorem, it remains to show that $H_{\min}(Y|AY'B)_{\rho} \leq H_{\min}(X|B)_{\rho} - \log \frac{1}{c}$ holds. More precisely, we will show that

$$\begin{aligned} H_{\min}(Y|Y'AB)_{\rho} &= \max_{\sigma} \sup \{ \lambda \in \mathbb{R} : \rho_{AYY'B} \leq 2^{-\lambda} \mathbb{1}_Y \otimes \sigma_{AY'B} \} \\ &\leq \sup \{ \lambda \in \mathbb{R} : \rho_{XB} \leq 2^{-\lambda} c \mathbb{1}_X \otimes \sigma_B \} \\ &= H_{\min}(X|B)_{\rho} - \log \frac{1}{c}. \end{aligned} \quad (7.10)$$

In order to arrive at (7.10), we thus need to show that, for any $\sigma \in \mathcal{S}_{=}(\mathcal{H}_{AY'B})$, the following implication holds

$$\rho_{AYY'B} \leq 2^{-\lambda} \mathbb{1}_Y \otimes \sigma_{AY'B} \implies \rho_{XB} \leq 2^{-\lambda} c \mathbb{1}_X \otimes \sigma_B. \quad (7.11)$$

To show this, we apply the partial isometry $W := UV^{\dagger}$ followed by a partial trace over X' and A on both sides of the inequality on the left-hand side. This implies

$$2^{\lambda} \rho_{XB} \leq \text{tr}_{X'A}(W(\mathbb{1}_Y \otimes \sigma_{AY'B})W^{\dagger}). \quad (7.12)$$

7.2 Traditional Formulation

Moreover, substituting the definition of W , we find that the trace term on the rhs. evaluates to

$$\text{rhs.} = \sum_{x,y} |x\rangle\langle x| \otimes \langle y| \text{tr}_A(\sqrt{M_x}\sqrt{N_y}\sigma_{AY'B}\sqrt{N_y}\sqrt{M_x})|y\rangle. \quad (7.13)$$

Lemma A.1, in particular Eq. (A.1), now establishes that

$$\text{tr}_A(\sqrt{M_x}\sqrt{N_y}\sigma_{AY'B}\sqrt{N_y}\sqrt{M_x}) \leq \|\sqrt{M_x}\sqrt{N_y}\|_\infty^2 \sigma_{Y'B} \leq c \cdot \sigma_{Y'B}.$$

Combining this with (7.13) and (7.12) results in the inequality

$$2^\lambda \rho_{XB} \leq 2^{-\lambda} c \sum_{x,y} |x\rangle\langle x| \otimes \langle y|\sigma_{BY'}|y\rangle = 2^{-\lambda} c \mathbb{1}_X \otimes \sigma_B.$$

This establishes (7.11) and concludes the proof. \square

7.2.3 Modeling the Measurement Basis Explicitly

An alternative formulation of the uncertainty principle requires an additional random variable, Θ , which determines the choice of measurement on the A system. Consider, for example, the setup of the previous section, where the choice is between two incompatible measurements, \mathcal{X} and \mathcal{Y} , which we assume both have the same number of different outcomes. In this case the random experiment of picking $\vartheta \in \Theta = \{0,1\}$ determines the binary choice of measurement. More specifically, say that $\vartheta = 0$ leads to a measurement of \mathcal{X} and $\vartheta = 1$ leads to a measurement of \mathcal{Y} . The measurement outcome, in either case, is stored in a classical register Z .

If Θ is uniform and independent of the state ρ before measurement, we find

$$H_{\min}(Z|B\Theta) + H_{\max}(Z|C\Theta) \geq \log \frac{1}{c(\mathcal{X},\mathcal{Y})}. \quad (7.14)$$

To see this, note that (7.8) implies

$$2^{-H_{\min}(X|B)} \leq c 2^{H_{\max}(Y|B)} \quad \text{and} \quad 2^{-H_{\min}(Y|B)} \leq c 2^{H_{\max}(X|B)}.$$

Taking the convex sum with equal weight $\frac{1}{2}$ of these two inequalities leads to (7.14), where we used the fact that the min- and max-entropies with classical side information can be expressed as averages (cf. Proposition 4.6).

A generalization of this type of uncertainty relation to smooth entropies will be discussed below in Corollary 7.4.

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

7.3 The Generalized Uncertainty Relation

7.3.1 Effective Overlap

The overlap $c(\mathcal{X}, \mathcal{Y})$ used in Theorem 7.1 is a function of the POVM elements of the two measurements under consideration and independent of the state prior to measurement. This is often desirable because this state might be unknown, or, in a cryptographic setting, prepared by an adversary. However, we will see that in some situations partial knowledge about the state before measurement can be used to improve the bound on the uncertainty. What follows is thus a generalized uncertainty relation of the form of Theorem 7.1 that introduces a trade-off between information about the marginal state before measurement and tightness of the uncertainty relation. Specifically, we consider the *effective overlap* of a *measurement setup*, denoted c^* , which describes the overlap of two measurements on a given *marginal* state.

Definition 7.2 (Effective Overlap). *Let $\rho_A \in \mathcal{S}_{\leq}(\mathcal{H}_A)$ be a state and let $\mathcal{X} = \{M_x\}$, $\mathcal{Y} = \{N_y\}$ be two POVMs on \mathcal{H}_A . Then, we call the triple $\{\rho_A, \mathcal{X}, \mathcal{Y}\}$ a measurement setup. The effective overlap of this measurement setup is defined as*

$$c^*(\rho_A, \mathcal{X}, \mathcal{Y}) := \min_{\mathcal{K}} \left\{ \sum_k \text{tr}(P^k \rho) \max_x \left\| P^k \sum_y N_y M_x N_y \right\|_{\infty} \right\}$$

where the minimum is taken over all projective measurements $\mathcal{K} = \{P^k\}$ on \mathcal{H}_A that commute with both \mathcal{X} and \mathcal{Y} .¹

In the following sections, we will show that an UCR also holds for this definition of effective overlap.

As a first example of the usefulness of such a generalized UCR, consider the scenario where we apply one of two projective measurements, either in the basis $\{|0\rangle, |1\rangle, |\perp\rangle\}$ or $\{|+\rangle, |-\rangle, |\perp\rangle\}$ on a state ρ which has the property that ‘ \perp ’ is measured with probability at most η .² A direct application of the state-independent uncertainty relation (Theorem 7.1) to this setup will not lead to the desired results as the overlap of the two bases is trivially $c = 1$. Still, our intuitive understanding of this situation tells us that the uncertainty about the measurement outcome is high as long as η is small.

In fact, the effective overlap of this setup satisfies $c_* \leq (1 - \eta)^{\frac{1}{2}} + \eta$. This formula can be interpreted as follows: with probability $1 - \eta$ we are in the subspace spanned by $|0\rangle$ and $|1\rangle$, where the overlap is $\frac{1}{2}$, and with probability η we measure \perp and have full overlap. To prove this upper bound,

¹The property that two measurements \mathcal{X} and \mathcal{K} commute is equivalent to the condition $M_x P^k = P^k M_x$ for all x and k .

²The diagonal states $|\pm\rangle$ are defined as $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$.

7.3 The Generalized Uncertainty Relation

simply choose the projective measurement $\mathcal{K} = \{|0\rangle\langle 0| + |1\rangle\langle 1|, |\perp\rangle\langle \perp|\}$ in Definition 7.2. Hence, while Theorem 7.1 thus only provides a trivial bound for this example, an uncertainty relation in terms of the effective overlap would give the expected bound.

We can always get a state-independent bound on the overlap by the choice $\mathcal{K} = \{\mathbb{1}_A\}$, that is

$$c^* \leq \max_x \left\| \sum_y N_y M_x N_y \right\|_\infty.$$

Furthermore, note that for projective measurements \mathcal{X} and \mathcal{Y} , the rhs. can be simplified to $\max_{x,y} |\langle x|y \rangle|^2$ agreement with the usual overlap in Definition 7.1. In general, it is conjectured that $c^* \leq c$, i.e. that

$$\max_x \left\| \sum_y N_y M_x N_y \right\|_\infty \leq \max_{x,y} \left\| \sqrt{N_y} M_x \sqrt{N_y} \right\|_\infty = \max_{x,y} \left\| \sqrt{M_x} \sqrt{N_y} \right\|_\infty^2. \quad (7.15)$$

Finally, if the measurements on the A system have binary outcomes, it is possible to upper bound the effective overlap (and, thus, lower bound the uncertainty) by the maximal CHSH Bell violation [CHSH69] that can be observed using this measurement setup on A with an arbitrary second party (i.e. on an arbitrary extension of the state and using arbitrary measurements by the other party.) This establishes an analytic relation between two fundamental concepts in quantum theory, Bell non-locality and uncertainty relations. We refer the interested reader to [HT11], where this relation is discussed in detail.

7.3.2 The Generalized Uncertainty Relation

We now consider a theorem that gives a very general formulation of the uncertainty principle for smooth entropies. It gives a lower bound on the uncertainty — in terms of smooth min- and max-entropies — about the outcome of two (incompatible) measurements, X and Z , conditioned on quantum side information and the result of an additional, projective measurement, K , that was done on the state before measuring X and Z .

More formally, we consider an arbitrary tripartite quantum state, ρ_{ABC} , two POVMs on A , $\mathcal{X} = \{M_x\}$ and $\mathcal{Y} = \{N_y\}$, as well as a projective measurement, $\mathcal{K} = \{P_k\}$. The post-measurement states when X and Y are measured after K , respectively, are thus

$$\rho_{XKB} = \sum_{x,k} |x\rangle\langle x| \otimes |k\rangle\langle k| \otimes \text{tr}_{AC}(\sqrt{M_x} P^k \rho_{ABC} P^k \sqrt{M_x}) \quad \text{and} \quad (7.16)$$

$$\rho_{YKC} = \sum_{y,k} |y\rangle\langle y| \otimes |k\rangle\langle k| \otimes \text{tr}_{AB}(\sqrt{N_y} P^k \rho_{ABC} P^k \sqrt{N_y}). \quad (7.17)$$

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

The following theorem generalizes previously known uncertainty relations for the smooth min- and max-entropies in a tripartite setting.¹ The UCRs discussed in Section 7.4 are corollaries of this relation.

Theorem 7.2 (Generalized UCR). *Let $\varepsilon \geq 0$, $\bar{\varepsilon} > 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$. Moreover, let $\mathcal{X} = \{M_x\}$, $\mathcal{Y} = \{N_y\}$ be POVMs on \mathcal{H}_A and $\mathcal{K} = \{P^k\}$ a projective measurement on \mathcal{H}_A . Then, the post-measurement states (7.16) and (7.17) satisfy*

$$H_{\min}^{\varepsilon}(X|KB)_{\rho} + H_{\max}^{\varepsilon}(Y|KC)_{\rho} \geq \log \frac{1}{c_{\mathcal{K}}} \quad \text{and} \quad (7.18)$$

$$H_{\min}^{2\varepsilon+\bar{\varepsilon}}(X|KB)_{\rho} + H_{\max}^{\varepsilon}(Y|KC)_{\rho} \geq \log \frac{1}{c_{\mathcal{K}}^*} - \log \frac{2}{\bar{\varepsilon}^2}. \quad (7.19)$$

where the \mathcal{K} -overlap, $c_{\mathcal{K}}$, and the effective \mathcal{K} -overlap, $c_{\mathcal{K}}^*$, are given as

$$c_{\mathcal{K}} := \max_{k,x,y} \left\| \sqrt{M_x} P^k \sqrt{N_y} \right\|_{\infty} \quad \text{and} \\ c_{\mathcal{K}}^* := \sum_k \text{tr}(P^k \rho) \max_x \left\| \sum_y P^k N_y P^k M_x P^k N_y P^k \right\|_{\infty}.$$

Proof. We prove the statement for pure ρ_{ABC} . Its generalization to mixed states then trivially follows from data processing inequalities for the smooth entropies (cf. Theorem 5.7). More specifically, assuming the result for pure states, we consider a purification ρ_{ABCD} of ρ_{ABC} , for which the theorem holds with the substitution $C \rightarrow CD$ and then take the partial trace over D . As this cannot decrease the smooth max-entropy, the generalization follows.

We consider the Stinespring dilation (cf. Lemma 2.4) of the joint measurement of \mathcal{X} and \mathcal{K} , denoted U , which coherently stores the measurement outcome of \mathcal{X} in registers X and X' and the measurement outcome of \mathcal{K} in K and K' . The isometry $U : \mathcal{H}_A \rightarrow \mathcal{H}_{AXX'KK'}$ is given by

$$U := \sum_{x,k} |x\rangle_X \otimes |x\rangle_{X'} \otimes |k\rangle_K \otimes |k\rangle_{K'} \otimes \sqrt{M_x} P^k.$$

Similarly, we introduce the Stinespring dilation of the joint measurement of \mathcal{Y} and \mathcal{K} as

$$V := \sum_{y,k} |y\rangle_Y \otimes |y\rangle_{Y'} \otimes |k\rangle_K \otimes |k\rangle_{K'} \otimes \sqrt{N_y} P^k.$$

We will also need the partial isometry $W := UV^{\dagger}$, which, using $P^k P^{k'} = \delta_{kk'} P^k$, evaluates to

$$W = \sum_{x,y,k} |x\rangle_X \langle y| \otimes |x\rangle_{X'} \langle y| \otimes |k\rangle_K \langle k| \otimes |k\rangle_{K'} \langle k| \otimes \sqrt{M_x} P^k \sqrt{N_y}. \quad (7.20)$$

¹The author is tempted — but will resist — to call it the mother uncertainty relation.

7.3 The Generalized Uncertainty Relation

These isometries allow us to introduce the states $\rho_{AXX'KK'BC} = U\rho_{ABC}U^\dagger$ and $\rho_{AYY'KK'BC} = V\rho_{ABC}V^\dagger$, whose marginals are the post-measurement states ρ_{XKB} and ρ_{YKC} of (7.16) and (7.17), respectively.

The proof now proceeds in several steps. First, we reformulate the statement of the theorem in terms of smooth min-entropies using the duality relation. Then, in order to show (7.18), we use techniques similar to the ones discussed in Section 7.2.2 to extract the \mathcal{K} -overlap. To show (7.19), we first use Lemma A.6 to find an upper bound on one of the entropies in terms of a relative entropy that is conditioned on ρ . The properties of ρ can then be used to extract the effective \mathcal{K} -overlap.

Due to the duality between smooth min- and max-entropy (cf. Theorem 5.4), the first statement of the theorem is equivalent to

$$H_{\min}^\varepsilon(X|KB)_\rho \geq H_{\min}^\varepsilon(Y|AY'K'B)_\rho + \log \frac{1}{c_{\mathcal{K}}}. \quad (7.21)$$

Let $\bar{\rho}$ be a state that maximizes the smooth min-entropy on the rhs. Then, there exists a σ such that

$$\lambda = S_{\min}(\bar{\rho}_{AYY'K'B} \| \mathbb{1}_Y \otimes \sigma_{AY'K'B}) = H_{\min}^\varepsilon(Y|AY'K'B)_\rho. \quad (7.22)$$

The state $\bar{\rho}$ can be chosen classical on K' due to Proposition 5.8. Furthermore, note that the minimum distance purification of $\bar{\rho}$, since its marginal is classical on K' , also inherits the coherence between K and K' from ρ . This follows from

$$P(\Pi\bar{\rho}\Pi, \rho) = P(\Pi\bar{\rho}\Pi, \Pi\rho\Pi) \leq P(\bar{\rho}, \rho),$$

where $\Pi = \sum_k |k\rangle\langle k|_K \otimes |k\rangle\langle k|_{K'}$ is the projector onto coherent superpositions of K and K' . We also used that $\text{tr}_K(\Pi\bar{\rho}\Pi) = \mathcal{M}[\bar{\rho}] = \bar{\rho}$, where \mathcal{M} measures the K' system; thus, the projection does not change the marginal state. We also need another extension of $\bar{\rho}$ to K , which we define via the TP-CPM $\mathcal{E} : \rho \mapsto \sum_k |k\rangle\langle k|_{K'} \rho |k\rangle\langle k|_{K'} \otimes |k\rangle\langle k|_K$. This map measures K' and creates a (classical) copy of it in K . Since $\bar{\rho}_{AYY'K'B}$ is classical on K , it is easy to see that $\bar{\tau}_{AYY'KK'B} = \mathcal{E}[\bar{\rho}_{AYY'K'B}]$ is an extension of $\bar{\rho}_{AYY'K'B}$. Furthermore, we can equivalently write $\bar{\tau}_{AYY'KK'B} = \Pi(\bar{\rho}_{AYY'K'B} \otimes \mathbb{1}_K)\Pi$ and note that the two extensions agree on the diagonals, i.e.

$$\forall k : \langle kk | \bar{\tau}_{AYY'KK'B} | kk \rangle = \langle kk | \bar{\rho}_{AYY'K'B} | kk \rangle. \quad (7.23)$$

From the definition of $S_{\min}(\bar{\rho}_{AYY'K'B} \| \mathbb{1}_Y \otimes \sigma_{AY'K'B})$, we get

$$\bar{\rho}_{AYY'K'B} \leq 2^{-\lambda} \mathbb{1}_Y \otimes \sigma_{AY'K'B}. \quad (7.24)$$

Taking the tensor product with $\mathbb{1}_K$ on both sides of (7.24) leads to

$$\bar{\rho}_{AYY'K'B} \otimes \mathbb{1}_K \leq 2^{-\lambda} \mathbb{1}_{YK} \otimes \sigma_{AY'K'B}.$$

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

We conjugate this inequality with W and take the partial trace over A , Y' and K' to get

$$\begin{aligned}\bar{\tau}_{XKB} &= \text{tr}_{AX'K'}(W(\bar{\rho}_{AYY'K'B} \otimes \mathbb{1}_K)W^\dagger) \\ &\leq 2^{-\lambda} \text{tr}_{AX'K'}(W(\mathbb{1}_{YK} \otimes \sigma_{AY'K'B})W^\dagger). \end{aligned} \quad (7.25)$$

We first evaluate the trace term on the rhs. of (7.25) using the expression for W in (7.20) and Lemma A.1 as in Eq. (A.1). We get

$$\begin{aligned}\text{tr}_{AX'K'}(W(\mathbb{1}_{YK} \otimes \sigma_{AY'K'B})W^\dagger) &= \sum_{x,y,k} |x\rangle\langle x| \otimes |k\rangle\langle k| \otimes \langle yk| \text{tr}_A(\sqrt{M_x}P^k\sqrt{N_y}\sigma_{AY'K'B}\sqrt{N_y}P^k\sqrt{M_x})|yk\rangle \\ &\leq \mathbb{1}_X \otimes \sum_k |k\rangle\langle k| \otimes \max_{x,y} \left\| \sqrt{N_y}P^k\sqrt{M_x} \right\|_\infty^2 \langle k|\sigma_{KB}|k\rangle \\ &= \mathbb{1}_X \otimes \tilde{\omega}_{KB}.\end{aligned}$$

The last equality defines the operator $\tilde{\omega}$. Note that $\text{tr}(\tilde{\omega}_{KB}) \leq c_K$; hence, using $\omega_{KB} = \tilde{\omega}_{KB}/c_K \in \mathcal{S}_\leq(\mathcal{H}_{KB})$ and (7.25), we find a lower bound on $S_{\min}(\bar{\tau}_{XKB} \parallel \mathbb{1}_X \otimes \omega_{KB})$ in terms of λ and c_K .

$$S_{\min}(\bar{\tau}_{XKB} \parallel \mathbb{1}_X \otimes \omega_{KB}) \geq \lambda + \log \frac{1}{c_K} = H_{\min}^\varepsilon(Y|AY'K'B)_\rho + \log \frac{1}{c_K}, \quad (7.26)$$

where we substituted (7.22) for λ . The state $\bar{\tau}_{XKB}$ is sub-normalized. Furthermore, its purified distance from ρ_{XKB} is bounded by

$$P(\bar{\tau}_{XKB}, \rho_{XKB}) = P(\bar{\rho}_{XKB}, \rho_{XKB}) \leq \varepsilon,$$

where we used that $\bar{\tau}_{XKB} = \bar{\rho}_{XKB}$. This follows from the property that the two states agree on the diagonals of $|kk\rangle$ — in the sense of (7.23) — together with the expression

$$\begin{aligned}\bar{\tau}_{XKB} &= \text{tr}_{AX'K'}(W\bar{\tau}_{AYY'KK'B}W^\dagger) = \sum_{k,x,y,y'} |x\rangle\langle x| \otimes |k\rangle\langle k| \otimes \\ &\text{tr}_A\left(\sqrt{M_x}P^k\sqrt{N_y}\langle yy'kk|\bar{\tau}_{AYY'KK'B}|y'y'kk\rangle\sqrt{N_{y'}}P^k\sqrt{M_x}\right),\end{aligned}$$

and the respective expression for $\bar{\rho}_{XKB}$. Hence, using the definition of the smooth min-entropy, we get

$$H_{\min}^\varepsilon(X|KB)_\rho \geq S_{\min}(\bar{\tau}_{XKB} \parallel \mathbb{1}_X \otimes \omega_{KB}),$$

which, substituted into (7.26), concludes the proof of (7.18).

It remains to show the second statement of the theorem. Again, due to the duality between smooth min- and max-entropy this is equivalent to

$$H_{\min}^{2\varepsilon+\bar{\varepsilon}}(X|KB)_\rho \geq H_{\min}^\varepsilon(Y|AY'K'B)_\rho + \log \frac{1}{c_K^*} - \log \frac{2}{\bar{\varepsilon}^2}. \quad (7.27)$$

7.3 The Generalized Uncertainty Relation

Applying Lemma A.6, we define the state $\tilde{\rho} \in \mathcal{B}^{2\varepsilon+\bar{\varepsilon}}(\rho)$ such that the following holds:

$$S_{\min}(\tilde{\rho}_{AYY'K'B} \| \mathbb{1}_Y \otimes \rho_{AY'K'B}) \geq H_{\min}^{\varepsilon}(Y|AY'K'B)_{\rho} - \log \frac{2}{\bar{\varepsilon}^2}.$$

We use the monotonicity of S_{\min} under TP-CPMs to measure the K system, i.e. we apply the map $\mathcal{M} : \tau \mapsto \sum_k |k\rangle\langle k|_{K'} \tau |k\rangle\langle k|_{K'}$ to both states in S_{\min} above. This will have no effect on $\rho_{AY'K'B}$, which is classical on K' by definition. Using the state $\bar{\rho}_{AYY'K'B} = \mathcal{M}[\tilde{\rho}_{AYY'K'B}]$, we have

$$\lambda = S_{\min}(\bar{\rho}_{AYY'K'B} \| \mathbb{1}_Y \otimes \rho_{AY'K'B}) \geq H_{\min}^{\varepsilon}(Y|AY'K'B)_{\rho} - \log \frac{2}{\bar{\varepsilon}^2}. \quad (7.28)$$

Moreover, the purified distance satisfies

$$P(\bar{\rho}, \rho) = P(\mathcal{M}[\tilde{\rho}], \mathcal{M}[\rho]) \leq P(\tilde{\rho}, \rho) \leq 2\varepsilon + \bar{\varepsilon}.$$

From the definition of $S_{\min}(\bar{\rho}_{AYY'K'B} \| \mathbb{1}_Y \otimes \rho_{AY'K'B})$, we get

$$\bar{\rho}_{AYY'K'B} \leq 2^{-\lambda} \mathbb{1}_Y \otimes \rho_{AY'K'B}, \quad (7.29)$$

where we employed the marginal state

$$\rho_{AY'K'B} = \text{tr}_{YK'}(V\rho_{AB}V^{\dagger}) = \sum_{y,k} \sqrt{N_y} P^k \rho_{AB} P^k \sqrt{N_y} \otimes |k\rangle\langle k| \otimes |y\rangle\langle y|. \quad (7.30)$$

The remainder follows the proof of the first statement closely. However, we will take advantage of the form of the marginal state in (7.30) to extract the effective \mathcal{K} -overlap. Taking the tensor product with $\mathbb{1}_K$ on both sides of (7.29) leads to

$$\bar{\rho}_{AYY'K'B} \otimes \mathbb{1}_K \leq 2^{-\lambda} \mathbb{1}_{YK} \otimes \rho_{AY'K'B}.$$

We conjugate this inequality with W and take the partial trace over A, Y' and K' to get

$$\begin{aligned} \bar{\tau}_{XKB} &= \text{tr}_{AX'K'}(W(\bar{\rho}_{AYY'K'B} \otimes \mathbb{1}_K)W^{\dagger}) \\ &\leq 2^{-\lambda} \text{tr}_{AX'K'}(W(\mathbb{1}_{YK} \otimes \rho_{AY'K'B})W^{\dagger}). \end{aligned} \quad (7.31)$$

We again evaluate the trace term on the rhs. of (7.31) using the expressions for W and $\rho_{AY'K'B}$ in (7.20) and (7.30). We get

$$\begin{aligned} &\text{tr}_{AX'K'}(W(\mathbb{1}_{YK} \otimes \rho_{AY'K'B})W^{\dagger}) \\ &= \sum_{x,y,k} |x\rangle\langle x| \otimes |k\rangle\langle k| \otimes \langle yk| \text{tr}_A(\sqrt{M_x} P^k \sqrt{N_y} \rho_{AY'K'B} \sqrt{N_y} P^k \sqrt{M_x}) |yk\rangle \\ &= \sum_x |x\rangle\langle x| \otimes \sum_k |k\rangle\langle k| \otimes \text{tr}_A\left(\sum_y \sqrt{M_x} P^k N_y P^k \rho_{AB} P^k N_y P^k \sqrt{M_x}\right) \\ &\leq \mathbb{1}_X \otimes \sum_k |k\rangle\langle k| \otimes \max_x \left\| \sum_y P^k N_y P^k M_x P^k N_y P^k \right\|_{\infty} \text{tr}_A(P^k \rho_{AB}) \\ &= \mathbb{1}_X \otimes \tilde{\omega}_{KB}. \end{aligned} \quad (7.32)$$

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

We used Lemma A.1 to arrive at (7.32). The last equality defines the operator $\tilde{\omega}$. Note that $\text{tr}(\tilde{\omega}_{KB}) = c_{\mathcal{K}}^*$; hence, we choose $\omega_{KB} = \tilde{\omega}_{KB}/c_{\mathcal{K}}^* \in \mathcal{S}_=(\mathcal{H}_{KB})$ and (7.31) and find a lower bound on $S_{\min}(\bar{\tau}_{XKB} \parallel \omega_{KB})$ in terms of λ and $c_{\mathcal{K}}^*$.

$$S_{\min}(\bar{\tau}_{XKB} \parallel \omega_{KB}) \geq \lambda + \log \frac{1}{c_{\mathcal{K}}^*} \geq H_{\min}^{\varepsilon}(Y|AY'K'B)_{\rho} + \log \frac{1}{c_{\mathcal{K}}^*} - \log \frac{2}{\bar{\varepsilon}^2}, \quad (7.33)$$

where we substituted (7.28) for λ . We have

$$P(\bar{\tau}_{XKB}, \rho_{XKB}) = P(\bar{\rho}_{XKB}, \rho_{XKB}) \leq 2\varepsilon + \bar{\varepsilon},$$

Hence, using the definition of the smooth min-entropy, we get

$$H_{\min}^{\varepsilon}(X|KB)_{\rho} \geq S_{\min}(\bar{\tau}_{XKB} \parallel \mathbb{1}_X \otimes \sigma_{KB}),$$

which, substituted into (7.33), concludes the proof of the second statement and the theorem. \square

7.4 Miscellaneous Uncertainty Relations

This section contains a collection of useful corollaries of Theorem 7.2, including the results discussed in the introduction of this chapter.

7.4.1 Commuting Measurements

A specialization of the generalized UCR which is of particular interest concerns the case when the two measurements \mathcal{X} and \mathcal{Y} both commute with \mathcal{K} . In this case, the marginal states of (7.16) and (7.17) when K is traced out correspond to the post-measurement states when only \mathcal{X} and \mathcal{Y} are measured, Eqs. (7.6) and (7.7). Formally,

$$\begin{aligned} \text{tr}_K(\rho_{XKB}) &= \sum_x |x\rangle\langle x| \otimes \text{tr}_{AC}(\sqrt{M_x} \rho_{ABC} \sqrt{M_x}) = \rho_{XB} \quad \text{and} \\ \text{tr}_K(\rho_{YKC}) &= \sum_y |y\rangle\langle y| \otimes \text{tr}_{AB}(\sqrt{N_y} \rho_{ABC} \sqrt{N_y}) = \rho_{YC}. \end{aligned}$$

They satisfy the following inequality.

Corollary 7.3. *Let $\varepsilon > 0$, $\varepsilon' \geq 0$, $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ and $\mathcal{X} = \{M_x\}$, $\mathcal{Y} = \{N_y\}$ two POVMs on \mathcal{H}_A . Then, the post-measurement states (7.6) and (7.7) satisfy*

$$H_{\min}^{2\varepsilon'+\varepsilon}(X|B)_{\rho} + H_{\max}^{\varepsilon'}(Y|C)_{\rho} \geq \log \frac{1}{c^*(\rho_A, \mathcal{X}, \mathcal{Y})} - \log \frac{2}{\varepsilon^2},$$

7.4 Miscellaneous Uncertainty Relations

Proof. Let \mathcal{K} be the measurement — commuting with \mathcal{X} and \mathcal{Y} — that minimizes the effective overlap, $c^*(\rho_A, \mathcal{X}, \mathcal{Y})$, in Definition 7.2. The corollary now follows from Theorem 7.2 applied to \mathcal{K} and the data processing inequality of the smooth min- and max-entropy (cf. Theorem 5.7) applied for the partial trace over K . \square

7.4.2 Modeling the Measurement Basis Explicitly

The full power of Theorem 7.2 comes to bear when we consider the following scenario. Let Θ be a classical register storing the choice of measurement that will be performed on the system A . For this purpose, we decompose $\mathcal{H}_{A'} \cong \mathcal{H}_\Theta \otimes \mathcal{H}_A$ and consider a family of POVMs $\{\mathcal{Z}^\vartheta\}$ on \mathcal{H}_A that share the same output alphabet. For all ϑ , let $\mathcal{Z}^\vartheta = \{L_z^\vartheta\}$ be the POVM on \mathcal{H}_A that is performed if ϑ is measured on Θ . This process can be equivalently modeled as a POVM \mathcal{Z} on $\mathcal{H}_{\Theta A}$, i.e.

$$\mathcal{Z} := \left\{ \sum_{\vartheta} |\vartheta\rangle\langle\vartheta| \otimes L_z^\vartheta \right\}_z. \quad (7.34)$$

If $\rho_{\Theta ABC} = \sum_{\vartheta} |\vartheta\rangle\langle\vartheta|_\Theta \otimes \rho_{ABC}^\vartheta$ is an arbitrary state classical on Θ , the post-measurement state is

$$\rho_{\Theta ZBC} = \sum_{\vartheta, z} |\vartheta\rangle\langle\vartheta| \otimes |z\rangle\langle z| \otimes \text{tr}_A \left(\sqrt{L_z^\vartheta} \rho_{ABC}^\vartheta \sqrt{L_z^\vartheta} \right). \quad (7.35)$$

This leads to the following result, which, in contrast to the uncertainty relations discussed above, relates the smooth min- and max-entropies of the same post-measurement state.

Corollary 7.4. *Let $\varepsilon \geq 0$ and $\rho_{\Theta ABC} = \sum_{\vartheta} |\vartheta\rangle\langle\vartheta|_\Theta \otimes \rho_{ABC}^\vartheta \in \mathcal{S}_\leq(\mathcal{H}_{\Theta ABC})$ be classical on Θ and let $f : \Theta \rightarrow \Theta$ be a bijective function such that $\rho_{ABC}^\vartheta = \rho_{ABC}^{f(\vartheta)}$. Then, the post-measurement state (7.35) after measuring the POVM (7.34) satisfies*

$$H_{\min}^\varepsilon(Z|\Theta B)_\rho + H_{\max}^\varepsilon(Z|\Theta C)_\rho \geq \log \frac{1}{c_f},$$

where $c_f = \max_{\vartheta} c(\mathcal{Z}^\vartheta, \mathcal{Z}^{f(\vartheta)})$.

Proof. We consider the POVMs $\mathcal{X} = \mathcal{Z}$ and $\mathcal{Y} = \left\{ \sum_{\vartheta} |\vartheta\rangle\langle\vartheta| \otimes L_z^{f(\vartheta)} \right\}_z$ as well as the projective measurement $\mathcal{K} = \{|\vartheta\rangle\langle\vartheta|\}$, $k = \vartheta$. It is easy to verify that \mathcal{K} commutes with \mathcal{X} and \mathcal{Y} and that the \mathcal{K} -overlap of Theorem 7.2 evaluates to $c_{\mathcal{K}} = c_f$. Thus,

$$H_{\min}^\varepsilon(X|\Theta B)_\rho + H_{\max}^\varepsilon(Y|\Theta C)_\rho \geq \log \frac{1}{c_f}.$$

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

Now, the corollary follows from the observation that the post-measurement states are equivalent up to local isometries, $\rho_{\Theta XB} = \rho_{\Theta ZB}$ and $\rho_{\Theta YC} = F^\dagger \rho_{\Theta ZC} F$, where $F = \sum_{\vartheta} |f(\vartheta)\rangle\langle\vartheta|$ and we used the natural isometries $|i\rangle \mapsto |i\rangle$ between the Hilbert spaces $\mathcal{H}_Z \cong \mathcal{H}_X \cong \mathcal{H}_Y$. Clearly, the smooth entropies are invariant under local isometries (cf. Proposition 5.3). \square

Note that a similar result can be derived based on the effective overlap formulation of the uncertainty relation, resulting in an overlap of $c_f^* = \sum_{\vartheta} \text{tr}(\rho^{\vartheta}) c(\mathcal{Z}^{\vartheta}, \mathcal{Z}^{f(\vartheta)})$.

Finally, the uncertainty relation (7.14) of Section 7.2.3 is a special case of this corollary where $\rho_{\Theta ABC} = \frac{1}{2} \mathbb{1}_{\Theta} \otimes \rho_{ABC}$ and f is the bit flip.

7.4.3 The von Neumann Limit

Using the asymptotic equipartition property, we directly get an uncertainty relation for the von Neumann entropy as well.

Corollary 7.5. *Let $\rho \in \mathcal{S}_=(\mathcal{H}_{ABC})$ and $\mathcal{X} = \{M_x\}$, $\mathcal{Y} = \{N_y\}$ two POVMs on \mathcal{H}_A and $\mathcal{K} = \{P^k\}$ a projective measurement on \mathcal{H}_A . Then, the post-measurement states (7.16) and (7.17) satisfy*

$$H(X|KB)_{\rho} + H(Y|KC)_{\rho} \geq \log \frac{1}{c_{\mathcal{K}}^*},$$

where $c_{\mathcal{K}}$ is defined in Theorem 7.2. Furthermore, it holds that

$$H(X|B)_{\rho} + H(Y|C)_{\rho} \geq \log \frac{1}{c^*(\rho_A, \mathcal{X}, \mathcal{Y})}.$$

Proof. We prove the first statement, from which the second follows by the same considerations that led to Corollary 7.3.

Consider an n -fold tensor product Hilbert space $\mathcal{H}_{A^n B^n C^n} \cong \mathcal{H}_{ABC}^{\otimes n}$ for an arbitrary $n \in \mathbb{N}$. We consider i.i.d. product states of ρ_{ABC} on this space, i.e. the states $\rho_{A^n B^n C^n} = \rho_{ABC}^{\otimes n}$. Furthermore, we define i.i.d. product measurements $\mathcal{X}^n = \mathcal{X}^{\otimes n} := \{\bigotimes_i M_{x_i}\}_{x^n}$, where $x^n = (x_1, x_2, \dots, x_n)$ is a string of n measurement outcomes. Similarly, $\mathcal{Y}^n = \mathcal{Y}^{\otimes n}$ and $\mathcal{K}^n = \mathcal{K}^{\otimes n}$. Clearly, the post-measurement states of the n -fold measurement setup also have i.i.d. product form, i.e. $\rho_{X^n K^n B^n} = \rho_{XKB}^{\otimes n}$ and $\rho_{Y^n K^n C^n} = \rho_{YKC}^{\otimes n}$.

The effective overlap of the n -fold setup can now be calculated as

$$\begin{aligned} c_{\mathcal{K}^n}^* &= \sum_{k^n} \text{tr} \left(\bigotimes_i P^{k_i} \rho \right) \max_{x^n} \left\| \sum_{y^n} \bigotimes_i P^{k_i} N_{y_i} P^{k_i} M_{x_i} P^{k_i} N_{y_i} P^{k_i} \right\|_{\infty} \\ &= \sum_{k^n} \prod_i \text{tr}(P^{k_i} \rho) \max_{x_i} \left\| \sum_{y_i} P^{k_i} N_{y_i} P^{k_i} M_{x_i} P^{k_i} N_{y_i} P^{k_i} \right\|_{\infty} \\ &= \prod_i \sum_{k_i} \text{tr}(P^{k_i} \rho) \max_{x_i} \left\| \sum_{y_i} P^{k_i} N_{y_i} P^{k_i} M_{x_i} P^{k_i} N_{y_i} P^{k_i} \right\|_{\infty} \\ &= (c_{\mathcal{K}}^*)^n \end{aligned} \tag{7.36}$$

7.4 Miscellaneous Uncertainty Relations

The uncertainty relation for smooth min- and max-entropies, Theorem 7.2, now states that

$$H_{\min}^{\varepsilon+2\varepsilon'}(X^n|K^n B^n)_\rho + H_{\max}^{\varepsilon'}(Z^n|K^n B^n)_\rho \geq \log \frac{1}{c\mathcal{K}^n} - \log \frac{2}{\varepsilon^2}.$$

for any $\varepsilon, \varepsilon' > 0$ such that $\varepsilon + 2\varepsilon' < 1$. We divide this inequality by n and use expression (7.36) for the effective overlap above to get

$$\frac{1}{n} H_{\min}^{\varepsilon+2\varepsilon'}(X^n|K^n B^n)_\rho + \frac{1}{n} H_{\max}^{\varepsilon'}(Z^n|K^n B^n)_\rho \geq \log \frac{1}{c\mathcal{K}} - \frac{1}{n} \log \frac{2}{\varepsilon^2}.$$

Taking the limit $n \rightarrow \infty$ using the asymptotic equipartition property (Corollaries 6.6 and 6.7), we find the uncertainty relation for von Neumann entropies. \square

7.4.4 The Quantum Key Distribution Setup

A specific example of the setup of Section 7.4.2 is very relevant in the application of the uncertainty relation to quantum key distribution. (See, for example, Section 8.3.)

We use $\Theta^n \in \{0, 1\}^n$ as a basis choice for a measurement on an n -partite system A^n . The measurement we consider is a product measurement, where, on the i -th part of the system, the measurement $\mathcal{Z}^{\vartheta_i} = \{L_z^{\vartheta_i}\}_z$ is executed depending on the corresponding bit ϑ_i in ϑ^n . The full POVM corresponding to this measurement process is

$$\mathcal{Z} = \left\{ \sum_{\vartheta^n} |\vartheta^n\rangle\langle\vartheta^n| \otimes \bigotimes_{i=1}^n L_{z^n}^{\vartheta_i} \right\}_{z^n}. \quad (7.37)$$

The result is an n -bit string z^n , which is stored in a classical register Z^n . Hence, the post-measurement state—when Θ^n is uniform and independent of ρ —is given as

$$\rho_{\Theta^n Z^n BC} = \sum_{\vartheta^n} \frac{1}{2^n} |\vartheta^n\rangle\langle\vartheta^n| \otimes \text{tr}_{A^n} \left(\bigotimes_i \sqrt{L_{z_i}^{\vartheta_i}} \rho_{A^n BC} \bigotimes_i \sqrt{L_{z_i}^{\vartheta_i}} \right). \quad (7.38)$$

Corollary 7.6. *Let $\varepsilon \geq 0$, $\bar{\varepsilon} > 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_{BE})$. Moreover, let \mathcal{Z}^0 and \mathcal{Z}^1 be two POVMs on \mathcal{H}_A . Then, the post-measurement state (7.38) that is produced by measuring \mathcal{Z} in (7.37) in a uniform and independent basis, Θ^n , satisfies*

$$H_{\min}^{\varepsilon}(Z^n|\Theta^n B)_\rho + H_{\max}^{\varepsilon}(Z^n|\Theta^n C)_\rho \geq n \log \frac{1}{c(\mathcal{Z}^0, \mathcal{Z}^1)}.$$

Furthermore, if the marginal $\rho_{A^n} = \rho_A^{\otimes n}$ is of i.i.d. product form, we have

$$H_{\min}^{2\varepsilon+\bar{\varepsilon}}(Z^n|\Theta^n B)_\rho + H_{\max}^{\varepsilon}(Z^n|\Theta^n C)_\rho \geq n \log \frac{1}{c^*(\rho_A, \mathcal{Z}^0, \mathcal{Z}^1)} - \log \frac{2}{\bar{\varepsilon}^2}.$$

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

Proof. To prove the first statement, we start with Corollary 7.4 and the measurements described above. The function f is in this case the bit-flip on the whole string Θ and the condition $\rho_{A^n BC}^\vartheta = \rho_{A^n BC}^{f(\vartheta)}$ is trivially satisfied since Θ is independent of ρ by assumption. Finally, it is easy to verify that the overlap c_f evaluates to $c(\mathcal{Z}^0, \mathcal{Z}^1)^n$.

The second statement requires an additional, projective measurement \mathcal{K} on \mathcal{H}_A that commutes with both \mathcal{Z}^0 and \mathcal{Z}^1 . Let \mathcal{K} be the measurement that minimizes the effective overlap $c^*(\rho_A, \mathcal{Z}^0, \mathcal{Z}^1)$ in Definition 7.2. We measure \mathcal{K} on all n subsystems together with Θ^n resulting in an additional classical register K^n . Using Theorem 7.2, this leads to the relation

$$\begin{aligned} & H_{\min}^{2\varepsilon+\bar{\varepsilon}}(Z^n|\Theta^n K^n B)_\rho + H_{\max}^\varepsilon(Z^n|\Theta^n K^n C)_\rho \\ & \geq n \log \frac{1}{c_{\mathcal{K}}^*(\rho_A, \mathcal{Z}^0, \mathcal{Z}^1)} - \log \frac{2}{\bar{\varepsilon}^2}, \end{aligned}$$

where we used the arguments in the proof of Corollary 7.5 to simplify the expression for the effective overlap. The statement then follows after a partial trace over K^n on both entropies. \square

Note that the above results can be extended to the case where measurements and the marginal state are of general product but not i.i.d. form. In this case, the logarithm of the overlap on the rhs. is replaced by an average over the logarithmic overlap on all subsystems.

7.5 Bipartite Uncertainty Relations

We have argued that the natural generalizations of uncertainty relation without side information to uncertainty relations with quantum side information introduces two distinct observers. Here, applying the chain rules for the smooth entropies introduced in Section 5.6, we show a bipartite uncertainty relation. Note that such an uncertainty relation necessarily needs to have a term on the rhs. that characterizes the entanglement between the observer and the system prior to measurement. If entanglement is present, the bound has to be reduced accordingly. In particular, if two projective measurements are executed on a system that is fully entangled with the observers, there is no uncertainty on the measurement outcome.¹ Hence, we see a trade-off between entanglement and uncertainty in bipartite uncertainty relations — the more entanglement there is prior to measurement, the less uncertainty is created.

Here, we show that this picture is still not complete. In order to get a tight result for arbitrary projective measurements and POVMs, we also need to consider the entanglement that is left after measurement. This is

¹See, for example [Ber08], where this is discussed for an uncertainty relation in the von Neumann limit and for rank-1 projective measurements.

7.5 Bipartite Uncertainty Relations

due to the fact that general measurements (in contrast to rank-1 projective measurements) do not necessarily destroy all entanglement between the two parties.

This result might have applications in two party quantum cryptography, where uncertainty relations are used to bound the knowledge of one party about the measurement results of the other party. An example of such an applications of bipartite UCRs can be found in [BFW11].

We consider a bipartite system shared between A and B , where, as before, the system A is measured using one of two measurements \mathcal{X} or \mathcal{Y} . The entanglement before measurement is characterized by the smooth min-entropy of A given B while the entanglement after measurement is given by the smooth max-entropy of the system after measurement, denoted A' given B and the measurement outcome. More precisely, we start with an arbitrary state ρ_{AB} and consider post-measurement states

$$\rho_{A'XB} = \sum_x |x\rangle\langle x| \otimes \sqrt{M_x} \rho_{AB} \sqrt{M_x} \quad \text{and} \quad (7.39)$$

$$\rho_{A'YB} = \sum_y |y\rangle\langle y| \otimes \sqrt{N_y} \rho_{AB} \sqrt{N_y}. \quad (7.40)$$

Note that, in contrast to the entropic uncertainty relations discussed earlier, the system A' will appear explicitly in our statements. Since the state of this system is not unique for a given POVM, we simply chose the simplest measurement TP-CPM consistent with the POVM, which leads to Eqs. (7.39) and (7.40). This implies the following uncertainty relation.¹

Theorem 7.7. *Let $\varepsilon \geq 0$, $\bar{\varepsilon} > 0$, $\tilde{\varepsilon} \geq 0$, $\tilde{\varepsilon}' \geq 0$ and $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Moreover, let \mathcal{X} and \mathcal{Y} be two POVMs on \mathcal{H}_A . Then, the post-measurement states (7.39) and (7.40) satisfy*

$$\begin{aligned} H_{\min}^{\hat{\varepsilon}}(X|B)_{\rho} + H_{\max}^{\varepsilon}(Y|B)_{\rho} &\geq H_{\min}^{\tilde{\varepsilon}}(A|B)_{\rho} - H_{\max}^{\tilde{\varepsilon}'}(A'|YB)_{\rho} \\ &\quad + \log \frac{1}{c^*} - 4 \log \frac{2}{\tilde{\varepsilon}^2}, \end{aligned}$$

where $\hat{\varepsilon} = 7\bar{\varepsilon} + 6\tilde{\varepsilon} + 4\tilde{\varepsilon}' + 8\varepsilon$ and $c^* = c^*(\rho_A, \mathcal{X}, \mathcal{Y})$.

Proof. We start from Corollary 7.3, which implies

$$H_{\min}^{\hat{\varepsilon}}(X|B)_{\rho} \geq H_{\min}^{3\bar{\varepsilon}+3\tilde{\varepsilon}+2\tilde{\varepsilon}'+4\varepsilon}(Y|A'Y'B)_{\rho} + \log \frac{1}{c^*} - \log \frac{2}{\tilde{\varepsilon}^2}.$$

Now, we apply the chain rule (5.13) to the min-entropy on the rhs. with the substitutions $\varepsilon \leftarrow \bar{\varepsilon}$, $\varepsilon' \leftarrow \tilde{\varepsilon}$ and $\varepsilon'' \leftarrow \bar{\varepsilon} + \tilde{\varepsilon}' + 2\varepsilon$ as well as the systems $A \leftarrow Y$, $B \leftarrow A'Y'$ and $C \leftarrow B$. This leads to

$$H_{\min}^{\hat{\varepsilon}}(X|B)_{\rho} \geq H_{\min}^{\tilde{\varepsilon}}(A'YY'|B)_{\rho} - H_{\max}^{\bar{\varepsilon}+\tilde{\varepsilon}'+2\varepsilon}(A'Y|B)_{\rho} + \log \frac{1}{c^*} - 3 \log \frac{2}{\tilde{\varepsilon}^2},$$

¹Other uncertainty relations of the types discussed above can be made bipartite too. Here, we are interested in a simple example.

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

where we used that Y and Y' are interchangeable. Then, we apply the chain rule (5.12) to the max-entropy with the substitutions $\varepsilon \leftarrow \bar{\varepsilon}$, $\varepsilon' \leftarrow \bar{\varepsilon}'$ and $\varepsilon'' \leftarrow \varepsilon$ as well as the systems $A \leftarrow Y'$, $B \leftarrow A'$ and $C \leftarrow B$. This directly leads to the statement of the theorem when we note that

$$H_{\min}^{\bar{\varepsilon}}(A'YY'|B)_\rho = H_{\min}^{\bar{\varepsilon}}(A|B)_\rho \quad (7.41)$$

due to the covariance of the smooth entropies under isometries (cf. Proposition 5.3). \square

We are also interested in the von Neumann limit of this uncertainty relation. Using the AEP and the same techniques as in the proof of Corollary 7.5, we find

$$H(X|B)_\rho + H(Y|B)_\rho \geq \log \frac{1}{c^*} + H(A|B)_\rho - H(A'|YB)_\rho.$$

Due to the symmetry of the lhs. of this expression, we can replace the Y on the rhs. by an X ; thus, the lower bound is effectively a function of the minimum of these two entropies. In the following, we prefer to model the basis explicitly and use that $H(Z|\Theta B) = \frac{1}{2}H(X|B) + \frac{1}{2}H(Y|B)$ if the basis choice is uniform. We get¹

$$H(Z|\Theta B)_\rho \geq \frac{1}{2} \log \frac{1}{c} + \frac{1}{2} \left(H(A|B)_\rho - H(A'|Z\Theta B)_\rho \right).$$

Note also that if the measurement is projective and rank-1, the system A' will simply contain a copy of Z and the second entropy thus vanishes. This leads to the uncertainty relation of [BCC⁺10],

$$H(Z|\Theta B)_\rho \geq \frac{1}{2} \left(\log \frac{1}{c} + H(A|B)_\rho \right).$$

7.5.1 Chained Uncertainty Relations

Consider two consecutive applications of the uncertainty relation—first measure Z_1 of a system A_1 in the basis Θ_1 and then Z_2 on A_2 in the basis Θ_2 . We can now derive two different bounds on the total entropy produced by these operations. First, we consider the two measurements together to get

$$H(Z|\Theta B)_\rho \geq \log \frac{1}{c} + \frac{1}{2} \left(H(A|B)_\rho - H(A'|Z\Theta B)_\rho \right), \quad (7.42)$$

where $\Theta = \Theta_1\Theta_2$, $Z = Z_1Z_2$, $A = A_1A_2$ and $A' = A'_1A'_2$ contains the systems A_1 and A_2 after measurement. On the other hand, we may write down the uncertainty of the first measurement separately.

$$H(Z_1|\Theta_1 B)_\rho \geq \frac{1}{2} \log \frac{1}{c} + \frac{1}{2} \left(H(A|B)_\rho - H(A'_1A_2|Z_1\Theta_1 B)_\rho \right).$$

¹For the following arguments, we replaced the effective overlap by the overlap.

7.5 Bipartite Uncertainty Relations

Now, the second system is measured. Namely, we consider a measurement of Z_2 in the basis Θ_2 on the second part of the joint system $A'_1 A_2$. The observers holds the information gained in the first measurement, i.e. Θ_1 and Z_1 in addition to B .

$$H(Z_2|Z_1\Theta B)_\rho \geq \frac{1}{2} \log \frac{1}{c} + \frac{1}{2} \left(H(A'_1 A_2|Z_1\Theta_1 B)_\rho - H(A'|Z\Theta B)_\rho \right).$$

Adding the previous two inequalities leads to

$$H(Z_1|\Theta_1 B)_\rho + H(Z_2|Z_1\Theta B)_\rho \geq \log \frac{1}{c} + \frac{1}{2} \left(H(A|B)_\rho - H(A'|Z\Theta B)_\rho \right). \quad (7.43)$$

The two bounds, Eqs. (7.42) and (7.43), are equivalent since Z_1 is independent of Θ_2 and, thus, $H(Z_1|\Theta_1 B) = H(Z_1|\Theta B)$.

This shows that we can split the process of producing uncertainty into two individual steps without loosening the bound on the uncertainty. In this sense, the bipartite uncertainty relation can be considered tight.

7. UNCERTAINTY RELATIONS FOR SMOOTH ENTROPIES

Chapter 8

Applications

This chapter discusses three example applications of the smooth entropy framework and uses results from Chapter 3-7.

In Section 8.1, we consider *source compression*, which has been used as an example in the introductory remarks, and show how the characterization of this task in the *one-shot* setting allows us to retrieve direct and converse bounds on coding for finite block-lengths as well as strong converse statements. In Section 8.2, we consider *randomness extraction*, a task that is very important in cryptography and is naturally considered in the one-shot setting. Then, in Section 8.3, we combine the above results and show how, in conjunction with an uncertainty relation, they allow us to prove security of a *quantum key distribution* protocol.

8.1 Full Characterization of Source Compression

This section provides an example of how the *one-shot* characterization of an information theoretic task is sufficient to derive bounds on the resource usage for finite block lengths and in the *i.i.d. limit*.

We consider source compression with quantum *side information*, or, equivalently, information reconciliation. In particular, we give a *strong converse* to information reconciliation. We employ many results of this thesis, including properties of the *purified distance* from Chapter 3, the *data processing* inequalities of Chapter 4 and the *asymptotic equipartition property* of Chapter 6.

The bounds derived here clearly also hold in case the side information is classical or non-existent. Thus, we provide bounds for classical source compression tasks as well. From this viewpoint, the following sections prove Shannon's source coding theorem [Sha48], as well as its extension to side information (Slepian-Wolf [SW73]) and quantum side information [DW03].¹ We also prove a strong *converse* of source compression

¹The Slepian-Wolf setting with two separate encoders can be viewed as source com-

8. APPLICATIONS

with side information. (See also [Ooh94], where the error exponent of the Slepian-Wolf strong converse is investigated.) Related to this work, a strong converse for compression of quantum information [Sch95] was shown by Winter [Win99].

8.1.1 One-Shot Characterization

The one-shot results are extracted from a recent paper by Renes and Renner [RR10]. Given a classical-quantum state ρ_{ZB} , shared between two parties, Alice and Bob, how much information needs to be transmitted from Alice to Bob such that Bob can reconstruct Z with probability of error at most ε ? To investigate this question, we consider non-interactive (one-way) information reconciliation protocols from Alice to Bob. These consist of an encoding function, $e : \mathcal{Z} \rightarrow \mathcal{M}$, that creates a message M to be sent from Alice to Bob. Bob then uses a decoder, a POVM $\mathcal{D} = \{D_{z'}\}$ acting on the joint state of B and M that produces a classical estimate of Z in the register Z' ($\mathcal{H}_Z \cong \mathcal{H}_{Z'}$). A protocol is thus characterized by the tuple $\mathcal{P} = \{e, \mathcal{D}\}$. Note that this is the most general model for a non-interactive information reconciliation protocol. In particular, since we consider a fixed state ρ_{ZB} , it is sufficient to consider deterministic encoding strategies

The initial state ρ_{ZB} is of the form

$$\rho_{ZB} = \sum_z P_Z(z) |z\rangle\langle z|_Z \otimes \rho_B^z, \quad (8.1)$$

where $P_Z(z)$ is the distribution over the input alphabet Z and $\rho_B^z \in \mathcal{S}_+(\mathcal{H}_B)$ are quantum states on B . Applying a protocol $\{e, \mathcal{D}\}$ as described above to this state results in the final state

$$\rho_{ZZ'} = \sum_{z, z'} P_Z(z) \text{tr} \left(D_{z'}(\rho_B^z \otimes |e(z)\rangle\langle e(z)|_M) \right) |z\rangle\langle z|_Z \otimes |z'\rangle\langle z'|_{Z'},$$

or, equivalently, the joint probability distribution $P_{ZZ'}(z, z') = P_Z(z) \cdot \text{tr} (D_{z'}(\rho_B^z \otimes |e(z)\rangle\langle e(z)|))$. The error probability of this protocol is

$$p_{\text{err}}(\mathcal{P}, \rho_{ZB}) := 1 - \sum_z P_Z(z) \text{tr} \left(D_z(\rho_B^z \otimes |e(z)\rangle\langle e(z)|_M) \right).$$

The error probability is thus equivalent to the trace distance between $\rho_{ZZ'}$ and the state $\chi_{ZZ'} = \sum_z P_Z(z) |z\rangle\langle z|_Z \otimes |z\rangle\langle z|_{Z'}$ describing perfect correlation between Z and Z' .

We characterize information reconciliation of a state ρ_{ZB} from Z to B with the minimum message length (in bits), m^ε , required to achieve an

pression followed by source compression with side information.

8.1 Full Characterization of Source Compression

error probability of at most ε .

$$m^\varepsilon(Z|B)_\rho := \min\{m : \exists \mathcal{P} \text{ s.t. } \log d_M = m \text{ and } p_{\text{err}}(\mathcal{P}, \rho_{ZB}) \leq \varepsilon\}.$$

We now slightly extend a result from [RR10]. The converse bound found there is not strong enough for our purposes when ε gets close to 1. However, this can be fixed easily.

Theorem 8.1. *Let $\rho_{ZB} \in \mathcal{S}_\leq(\mathcal{H}_{ZB})$ be classical on Z and let $0 < \varepsilon < 1$. Then,*

$$H_{\max}^{\sqrt{2\varepsilon - \varepsilon^2}}(Z|B)_\rho \leq m^\varepsilon(Z|B)_\rho \leq H_{\max}^{\varepsilon_1}(Z|B)_\rho + 2 \log \frac{1}{\varepsilon_2} + 4,$$

for any $\varepsilon_1, \varepsilon_2$ s.t. $\varepsilon_1 + \varepsilon_2 = \varepsilon$.

Proof. The direct bound, $m^\varepsilon(Z|B)_\rho \leq H_{\max}^{\varepsilon_1}(Z|B)_\rho + 2 \log \frac{1}{\varepsilon_2} + 4$, is shown in [RR10]. To get a converse bound, we also follow their argument and note that $p_{\text{err}}(\mathcal{P}, \rho_{ZB}) \leq \varepsilon$ is equivalent to the condition $D(\rho_{ZZ'}, \chi_{ZZ'}) \leq \varepsilon$. This implies $P(\rho_{ZZ'}, \chi_{ZZ'}) \leq \sqrt{2\varepsilon - \varepsilon^2} = \varepsilon'$ according to Proposition 3.3. Hence,

$$H_{\max}^{\varepsilon'}(Z|Z')_\rho \leq H_{\max}(Z|Z')_\chi = 0 \quad (8.2)$$

by definition of the smooth max-entropy. The data-processing inequality (Theorem 5.7) then states that $H_{\max}^{\varepsilon'}(Z|MB)_\rho \leq 0$ before measurement. And finally we use the fact that conditioning on classical information M can at most reduce the smooth max-entropy by $\log d_M$ (cf. Proposition 5.10). This leads to the following inequality:

$$0 \geq H_{\max}^{\varepsilon'}(Z|MB)_\rho \geq H_{\max}^{\varepsilon'}(Z|B)_\rho - \log d_M.$$

Hence, all protocols with $p_{\text{err}}(\mathcal{P}, \rho_{ZB}) \leq \varepsilon$ have to satisfy the constraint $\log d_M \geq H_{\max}^{\varepsilon'}(Z|B)_\rho$. This results in the improved converse bound of the theorem. \square

8.1.2 Finite Block Lengths and Asymptotic Limits

The one-shot result in principle characterizes all data-reconciliation tasks. Comparing the one-shot analysis with asymptotic results, where m^ε converges to the *von Neumann entropy* [DW03], we might be interested to know how fast the one-shot result in terms of the smooth max-entropy converges to the von Neumann entropy.

For this purpose, we consider the task of encoding a block of length n of *i.i.d.* random variables Z about each of which Bob independently has quantum side information B . This is data reconciliation for the state

8. APPLICATIONS

$\rho_{Z^n B^n} = \rho_{ZB}^{\otimes n}$ and Theorem 8.1 applies. The minimum message length (per round) for this task, $\bar{m}^\varepsilon(Z|B) = \frac{1}{n}m^\varepsilon(Z^n|B^n)$, thus satisfies

$$\frac{1}{n}H_{\max}^{\sqrt{2\varepsilon-\varepsilon^2}}(Z^n|B^n)_\rho \leq \bar{m}^\varepsilon(Z|B)_\rho \leq \frac{1}{n}\left(H_{\max}^{\varepsilon/2}(Z^n|B^n)_\rho + 2\log\frac{1}{\varepsilon} + 6\right), \quad (8.3)$$

where we chose $\varepsilon_1 = \varepsilon_2 = \frac{\varepsilon}{2}$ for convenience of exposition. Next, we bound the smooth max-entropies using the asymptotic equipartition property in its entropic form.

Direct Bounds

We start with the upper bound in (8.3). For sufficiently large n , the AEP states that (cf. Theorem 6.4 and Corollary 6.5)

$$\bar{m}^\varepsilon(Z|B)_\rho \leq H(Z|B)_\rho + \frac{4\log v \sqrt{g(\frac{\varepsilon}{2})}}{\sqrt{n}} + \frac{2\log\frac{1}{\varepsilon} + 6}{n}, \quad (8.4)$$

where $g(\varepsilon) = -\log(1 - \sqrt{1 - \varepsilon^2})$ and $v = \sqrt{2^{-H_{\min}(Z|B)_\rho}} + \sqrt{2^{H_{\max}(Z|B)_\rho}} + 1$. Hence, in the asymptotic limit of large n , we have

$$\lim_{n \rightarrow \infty} \bar{m}^\varepsilon(Z|B)_\rho \leq H(Z|B)_\rho.$$

Namely, there exists an encoding strategy that transmits at most $H(Z|B)_\rho$ bits per round to Bob with an arbitrarily small positive probability of error ($\varepsilon > 0$). The converse question is whether we can do better than this.

Converse Bounds

To see that this is not possible, even for large error probabilities, we apply the AEP to the lower bound in (8.3). In the following, we use ε to denote the success probability, i.e. $1 - \varepsilon$ is the error probability. Using Proposition 5.5 and Remark 5.6, we find

$$\begin{aligned} \bar{m}^{1-\varepsilon}(Z|B)_\rho &\geq \frac{1}{n}H_{\max}^{\sqrt{1-\varepsilon^2}}(Z^n|B^n)_\rho \\ &\geq \frac{1}{n}\left(H_{\min}^{\tilde{\varepsilon}}(Z^n|B^n)_\rho - 2\log\frac{1}{1 - (\varepsilon\tilde{\varepsilon} + \sqrt{1-\varepsilon^2}\sqrt{1-\tilde{\varepsilon}^2})^2}\right). \end{aligned}$$

Here, the smoothing parameter of the min-entropy, $\tilde{\varepsilon}$, is restricted by the inequality $\arcsin(\tilde{\varepsilon}) + \arccos(\varepsilon) < \frac{\pi}{2}$. This holds if and only if $\tilde{\varepsilon} < \varepsilon$ and, for convenience of exposition, we choose $\tilde{\varepsilon} = \frac{\varepsilon}{2}$. This ensures that the logarithmic correction term is finite for any $0 < \varepsilon < 1$. For the fun of

8.1 Full Characterization of Source Compression

introducing a symmetry with Eq. (8.4) where there is none (at least not in the form of the explicit deviation terms), we further bound this as follows.¹

$$\bar{m}^{1-\varepsilon}(Z|B)_\rho \geq \frac{1}{n} \left(H_{\min}^{\varepsilon/2}(Z^n|B^n)_\rho - 2 \log \frac{1}{\varepsilon} - 6 \right),$$

Using the AEP, this time for the min-entropy, leads to

$$\bar{m}^{1-\varepsilon}(Z|B)_\rho \geq H(Z|B)_\rho - \frac{4 \log v \sqrt{g(\frac{\varepsilon}{2})}}{\sqrt{n}} - \frac{2 \log \frac{1}{\varepsilon} + 6}{n}. \quad (8.5)$$

Hence, in the asymptotic limit, for all $0 < \varepsilon < 1$, we find

$$\lim_{n \rightarrow \infty} \bar{m}^{1-\varepsilon}(Z|B)_\rho \geq H(Z|B)_\rho.$$

This is the converse of data-reconciliation. It states that, even for an arbitrarily small success probability ($\varepsilon > 0$), there does not exist a protocol that allows us to decode using less than $H(Z|B)_\rho$ transmitted bits per round.

8.1.3 Strong Converse

Finally, the non-asymptotic converse bound, Eq. (8.5) gives us the means to make this statement more precise. For example, we can ask how the maximal success probability, ε_{\max} , scales if we use a protocol that encodes only $H(Z|B)_\rho - \mu$ bits per round for large n . For this analysis, we neglect the term scaling reciprocally in n , i.e. we consider the approximate bound

$$\bar{m}^{1-\varepsilon}(Z|B)_\rho \gtrsim H(Z|B)_\rho - \frac{4 \log v \sqrt{g(\frac{\varepsilon}{2})}}{\sqrt{n}}.$$

This implies that

$$\begin{aligned} \varepsilon_{\max}(\mu) &\lesssim \sup \left\{ \varepsilon > 0 : \frac{4 \log v \sqrt{g(\frac{\varepsilon}{2})}}{\sqrt{n}} \geq \mu \right\} \\ &\leq \sqrt{2} \cdot 2^{-\frac{\mu^2 n}{2(4 \log v)^2}}. \end{aligned} \quad (8.6)$$

In order to find (8.6), we used the bound $g(\varepsilon/2) \leq \log \frac{8}{\varepsilon^2}$ (cf. Figure 6.2). This establishes that the maximal success probability drops exponentially in n and μ^2 if we transmit less than the Shannon limit, $H(Z|B)_\rho$ bits per round. The exponential drop of the success probability in n is called the strong converse.

¹We hope the reader does not mind us skipping the details of this simple procedure, which we have tested numerically and analytically.

8. APPLICATIONS

8.2 Randomness Extraction

This section discusses *randomness extraction*, the art of extracting uniform randomness from a biased source. More precisely, we consider a source that outputs a register Z about which there exists *side information* E —potentially quantum—and ask how much uniform randomness S , independent of E , can be extracted from Z . This primitive is of crucial importance in many cryptographic tasks, for example in quantum cryptography. There, we are interested to distill a secret key from a raw key that is partially correlated with a quantum eavesdropper.

The quality of the extracted randomness is measured using the trace distance; we consider the distance

$$\Delta(S|E)_\rho := \min_{\sigma_E} D(\rho_{SE}, \pi_S \otimes \sigma_E), \quad \text{where} \quad \pi_S = \frac{1}{d_S} \mathbb{1}_S$$

and the optimization is over all $\sigma_E \in \mathcal{S}_=(\mathcal{H}_E)$. Due to the operational interpretation of the trace distance as a *distinguishing advantage* (3.1), a small Δ implies that the extracted random variable cannot be distinguished from a uniform and independent random variable with probability more than $\frac{1}{2}(1 + \Delta)$. This viewpoint is at the root of the universally composable security framework [Can01, PW01] in cryptography, which was extended to the quantum setting in [Unr10] based on earlier work, for example in [Ren05].

We allow probabilistic protocols to extract this uniform randomness. These can be modeled by introducing an additional independent random seed, stored in a register F and then applying a (deterministic) function $f \in F$ on X to get S . Namely, we consider a joint state ρ_{ZEF} of the form

$$\rho_{ZEF} = \sum_z P_Z(z) |z\rangle\langle z|_Z \otimes \rho_E^z \otimes \sum_f P_F(f) |f\rangle\langle f|_F, \quad \rho_E^z \in \mathcal{S}_\leq(\mathcal{H}_E).$$

and require that $\sum P_F(f) \Delta(S|E)_{\rho^f}$ is small, where ρ_{SE}^f is the state produced when f is applied to the Z register of ρ_{ZE} .¹ A randomness extraction protocol, \mathcal{P} , thus consists of a *probability distribution* $P_F(f)$ on a *register* F and a family of functions $f \in F$ from Z to S .

The maximal number of bits of uniform randomness, ℓ^ε , that can be extracted from a state ρ_{ZE} is defined as

$$\ell^\varepsilon(Z|E)_\rho := \max \left\{ \ell : \exists \mathcal{P} \text{ s.t. } \log d_S = \ell \text{ and } \sum P_F(f) \Delta(S|E)_{\rho^f} \leq \varepsilon \right\}.$$

A protocol that satisfies $\sum P_F(f) \Delta(S|E)_{\rho^f} \leq \varepsilon$ is called ε -good.

¹This is satisfied if $\Delta(S|EF)_\rho$ is small, i.e. if the extracted randomness is independent of the seed F as well as E [TSSR11]. This is also called the strong extractor regime in classical cryptography.

8.2.1 Direct Bound

A particular class of protocols that can be used to extract uniform randomness are based on two-universal hashing (see Carter and Wegmann [CW79]). The classical Leftover Hash Lemma [Mc187, ILL89, IZ89] states the amount of extractable randomness is at least the min-entropy of Z given E . This construction was then extended to the quantum setting in [KMR05, RK05, Ren05]. In fact, since hashing is an entirely classical process, one may expect that the physical nature of the side information is irrelevant and that a purely classical treatment is sufficient. This is, however, not necessarily the case. For example, the output of certain extractor functions may be partially known if side information about their input is stored in a quantum device of a certain size, while the same output is almost uniform conditioned on any side information stored in a classical system of the same size.¹ For protocols based on two-universal hashing, the following *direct bound* holds [TSSR11].

$$\ell^\varepsilon(Z|E)_\rho \geq H_{\min}^{\varepsilon_1}(Z|E)_\rho - 2 \log \frac{1}{\varepsilon_2} + 1, \quad \text{where } \varepsilon = \varepsilon_1 + \varepsilon_2. \quad (8.7)$$

Other protocol families that extract the min-entropy against quantum adversaries² are based on almost two-universal hashing [TSSR11] or Trevisan extractors [DPVR09].

Note that the protocol families discussed above work on any state ρ_{ZE} with sufficiently high min-entropy, i.e. they do not take into account other properties of the state. Next, we will see that these protocols are essentially optimal.

8.2.2 Converse Bound

We prove a *converse bound* by contradiction. This formalizes an intuitive argument given in [TSSR11]. Assume for the sake of the argument that we have an ε -good protocol that extracts $\ell > H_{\min}^{\varepsilon'}(Z|E)_\rho$ bits of randomness, where $\varepsilon' = \sqrt{2\varepsilon - \varepsilon^2}$. Then, due to Proposition 5.11 we know that applying a function on Z cannot increase the smooth min-entropy, thus

$$\forall f \in F : H_{\min}^{\varepsilon'}(S|E)_{\rho^f} < \ell \quad \text{and, thus,} \quad \Delta(S|E)_{\rho^f} > \varepsilon. \quad (8.8)$$

The second statement of (8.8) follows from the following observation. The inequality $H_{\min}^{\varepsilon'}(S|E)_{\rho^f} < \ell$ implies that all states $\tilde{\rho}$ with $D(\tilde{\rho}_{SE}, \rho_{SE}^f) \leq \varepsilon$ (cf. Proposition 3.3) must necessarily satisfy $H_{\min}(S|E)_{\tilde{\rho}} < \ell$. In particular, these close states can thus not be of the form $\pi_S \otimes \sigma_E$, because such states have min-entropy ℓ . Hence, $\Delta(S|E)_{\rho^f} > \varepsilon$.

¹See [GKK⁺07] for a concrete example and [KR11] for a more general discussion).

²These families are considered mainly because they need a smaller seed or can be implemented more efficiently than two-universal hashing.

8. APPLICATIONS

However, (8.8) implies that $\sum P_F(f) \Delta(S|E)_{\rho^f} > \varepsilon$, contradicting our initial assumption that the protocol was ε -good. This implies the following converse bound

$$\ell^\varepsilon(Z|E)_\rho \leq H_{\min}^{\sqrt{2\varepsilon-\varepsilon^2}}(Z|E)_\rho. \quad (8.9)$$

Collecting (8.7) and (8.9), we have the following theorem

Theorem 8.2. *Let $\rho_{ZE} \in \mathcal{S}_\leq(\mathcal{H}_{ZE})$ be classical on Z and let $0 < \varepsilon < 1$. Then,*

$$H_{\min}^{\varepsilon_1}(Z|E)_\rho - 2 \log \frac{1}{\varepsilon_2} + 1 \leq \ell^\varepsilon(Z|E)_\rho \leq H_{\min}^{\sqrt{2\varepsilon-\varepsilon^2}}(Z|E)_\rho,$$

for any $\varepsilon_1, \varepsilon_2$ s.t. $\varepsilon_1 + \varepsilon_2 = \varepsilon$.

We have established that the extractable uniform and independent randomness is characterized by the smooth min-entropy. A similar analysis for finite block lengths and the *i.i.d. limit* as in Section 8.1 is thus possible. However, we omit it here since most applications consider the task of randomness extraction only in the *one-shot* setting. An example of such an application, quantum key distribution, follows in the next section.

8.3 Security in Quantum Key Distribution

The smooth entropy formalism was first introduced in the quantum setting by Renner [Ren05] in order to prove information theoretic security of *quantum key distribution* (QKD) protocols [BB84, Eke91] against adversaries restricted only by the laws of physics. Prior to this work, the security of QKD protocols was mostly analyzed in the limit of long keys and using questionable security definitions [KRBM07].

In this section, we show that the uncertainty relations for smooth entropies in Chapter 7 can be employed to provide a very concise and intuitive security proof for QKD. This is based on two recent papers [TR11] and [TLGR11]. The latter paper contains a thorough analysis of the statistical tests required to assert security of the final key that goes beyond the discussion here.

8.3.1 The Protocol

We consider the original BB84 quantum key distribution protocol [BB84]. Here, two legitimate parties, Alice and Bob, try to distill a shared secret key that is independent of all wiretapped information. To do this, Alice and Bob share a public quantum channel and a public, authenticated classical

8.3 Security in Quantum Key Distribution

channel. In the entanglement-based version of the BB84 protocol [BBM92], Alice now prepares n fully entangled qubit pairs of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{A_i} \otimes |0\rangle_{B_i} + |1\rangle_{A_i} \otimes |1\rangle_{B_i}), \quad (8.10)$$

where $i \in \{1, \dots, n\}$. She then sends the systems $B^n = B_1 \dots B_n$ over the public quantum channel to Bob.

Next, Alice chooses a basis for each qubit, $\Theta_i \in \{0, 1\}$, uniformly at random. She communicates the basis choices, $\Theta^n = \Theta_1 \dots \Theta_n$, to Bob over the classical channel and then measures her qubits A_i in the basis $\mathcal{X} = \{|0\rangle, |1\rangle\}$ if $\Theta_i = 0$ or $\mathcal{Y} = \{|+\rangle, |-\rangle\}$ if $\Theta_i = 1$.¹ The collected measurement outcomes, called raw key, are stored in a string $Z^n = Z_1 \dots Z_n$. Bob does a measurement on each of his systems depending on the value of Θ_i to produce a raw key \bar{Z}^n . This already concludes the quantum part of the protocol.

Next, Alice and Bob employ a one-way data reconciliation protocol that allows Bob to create an estimate \hat{Z}^n of Z^n using his raw key \bar{Z}^n . Finally, Alice invokes a randomness extraction protocol on her string Z to distill a key S and broadcasts the seed, F , over the classical channel in order to allow Bob to extract an estimate \hat{S} of S . The latter step is usually called privacy amplification [BBR88, BBCM95].

We note that, if the eavesdropper does not interfere and Bob simply measures his part of the entangled pair using the same bases as Alice, the resulting raw keys, Z^n and \bar{Z}^n , will be perfectly correlated. In this case, security of the key simply follows from the monogamy of entanglement (i.e. Bob is the preferred observer of Alice's system) and no information reconciliation is necessary. In the next section, we analyze this protocol in the presence of an eavesdropper and noise. (The effects of noise and wiretapping can in general not be distinguished; hence, we consequently assume that correlations are degraded due to wiretapping on the quantum channel.)

8.3.2 Security in the Finite Key Regime

Under what conditions will the key extracted by the above protocol be both secret (i.e. uniform and independent of the eavesdropper's information) and correct (i.e. $S = \hat{S}$)? We will see that quantum mechanics allows us to ascertain both secrecy and correctness from the correlations between Alice's measurement outcomes and Bob's side information about them. This is in contrast to classical theory, where, without further assumptions, we could only hope to determine correctness from such correlations. Quantum mechanics enables this due to the asymmetry between different

¹The diagonal basis is given by $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$.

8. APPLICATIONS

quantum mechanical observers, as discussed in the introduction of this thesis. If it can be established that one observer, Bob, is a preferred observer of a quantum system Alice holds, then this ensures that an eavesdropper's knowledge about the outcomes of measurements performed on that system is limited. Here, we show how this asymmetry can be verified using the entropic uncertainty relation of Chapter 7.

Since the eavesdropper is allowed to interact arbitrarily with the quantum communication sent from Alice to Bob, we do not make assumptions about the form of the resulting state $\rho_{A^n B^n E}$. In particular, the systems B_i that arrive at Bob's lab do not need to be qubits and E is an arbitrarily large system, held by the eavesdropper, that may be correlated with the systems $A^n = A_1 \dots A_n$ and B^n .

The uncertainty relation for smooth min- and max-entropies, in particular Corollary 7.6, can be applied to this setup. It states that

$$H_{\min}^\varepsilon(Z^n | \Theta^n E) + H_{\max}^\varepsilon(Z^n | \Theta^n B) \geq n \log \frac{1}{c}, \quad (8.11)$$

where c is the overlap of Alice's measurements (one one qubit). In the case of entanglement-based BB84 as described above, we have $c = \frac{1}{2}$ and, thus, the left-hand side of (8.11) evaluates to n . This means that the smooth min-entropy of the eavesdropper's information about Z^n right after Alice's measurements is lower bounded by n minus the max-entropy of Bob's correlations with Z^n . After Bob's measurement, using the data-processing inequality (cf. Theorem 5.7), we thus have

$$H_{\min}^\varepsilon(Z^n | \Theta^n B) \geq n - H_{\max}^\varepsilon(Z^n | \bar{Z}^n).$$

During information reconciliation, using a protocol as discussed in the direct part of Theorem 5.7, Alice sends a message M to Bob which satisfies

$$\log d_M \leq H_{\max}^{\varepsilon_c/2}(Z^n | \bar{Z}^n) + 2 \log \frac{1}{\varepsilon_c} + 6,$$

where ε_c is the required correctness. This ensures that the probability that Bob cannot correctly estimate Alice's string is at most ε_c . Note that this message might be wiretapped by the eavesdropper and stored in a register M' . However, using Proposition 5.10, we bound

$$\begin{aligned} H_{\min}^\varepsilon(Z^n | \Theta^n EM) &\geq H_{\min}^\varepsilon(Z^n | \Theta^n E) - \log d_M \\ &\geq n - H_{\max}^\varepsilon(Z^n | \bar{Z}^n) - H_{\max}^{\varepsilon_c/2}(Z^n | \bar{Z}^n) - 2 \log \frac{1}{\varepsilon_c} - 6. \end{aligned} \quad (8.12)$$

The direct part of Theorem 8.2 now states that we can extract at least ℓ bits of ε_s -secure key, where

$$\ell \geq H_{\min}^{\varepsilon_s/2}(Z^n | \Theta^n EM) - 2 \log \frac{1}{\varepsilon_s} - 1.$$

8.3 Security in Quantum Key Distribution

By the term ε_s -secure we mean that $\Delta(S|\Theta^n EMF)_\rho \leq \varepsilon_s$. This implies that the secret key is in particular independent of the seed F , which Alice needs to send to Bob and which could be wiretapped by the eavesdropper.

Using (8.12), we can express this bound entirely in terms of correlations between Alice and Bob, i.e.

$$\ell \geq n - H_{\max}^{\varepsilon_s/2}(Z^n|\bar{Z}^n) - H_{\max}^{\varepsilon_c/2}(Z^n|\bar{Z}^n) - 2 \log \frac{1}{\varepsilon_c} - 2 \log \frac{1}{\varepsilon_s} - 7.$$

This means that if the classical correlations between Alice's and Bob's measurement outcomes are sufficiently good — namely if the corresponding smooth max-entropies are small — we can safely extract a secret key using this protocol.

It thus remains to find a statistical test to determine upper bounds on the smooth max-entropies. This is usually done in the following way. Alice and Bob, after measuring, compare a random subset of their raw keys using the classical channel. Then, if the frequency of errors found in this sample is smaller than an agreed threshold value, they will proceed with the protocol and extract a key. In this case, it can be shown that the smooth max-entropy is indeed small, and the key thus secure. Otherwise, they abort and do not produce a key. (A detailed analysis of the statistical tests and the precise security statements that follow is beyond the scope of this thesis. Such an analysis can be found in [TLGR11], for example.)

8. APPLICATIONS

Chapter 9

Conclusions and Outlook

The goal of this thesis was to consolidate the smooth entropy framework for non-asymptotic information theory and to introduce important additions to the framework, including the entropic asymptotic equipartition property and various uncertainty relations. My sincere hope is that this work will provide a reference for researchers interested in non-asymptotic quantum information theory and quantum cryptography.

9.1 Applications of the Framework

The smooth entropy framework has already found a wide range of applications since its conception.

In cryptography, particularly quantum key distribution, the smooth entropy framework has become a standard tool to analyze security for finite keys [Ren05, SR08]. This analysis was simplified for some protocols thanks to the entropic uncertainty relation [TSSR11, TLGR11]. The smooth entropy formalism allows to investigate entropically secure encryption [FS08, DD10], and, together with results from randomness extraction, we use it to show that bit commitments cannot be expanded [WTHR11] in a quantum world. Furthermore, composable security in the bounded storage [WW08] and noisy storage [KWW09] models is analyzed using smooth entropies.

Decoupling of quantum systems [DBWR10] can be viewed as a fully quantum generalization of randomness extraction and is characterized by smooth entropies. The decoupling approach — a quantum generalization of random coding — leads to direct bounds for many information theoretic tasks. (Some of them are listed in [Dup09] and [DBWR10].) We have also shown that decoupling is possible using approximate two-designs [SDTR11], which suggests that decoupling can be achieved efficiently in nature.

The smooth entropy formalism has also been used to investigate various quantum channel capacities and converses [BD10, DHB11, DH11]. In

9. CONCLUSIONS AND OUTLOOK

particular, it leads to a conceptually simple proof of the quantum Reverse Shannon Theorem [BCR11].

In thermodynamics, smooth entropies have been used to quantify work extraction [DRRV11] and the work cost of erasure [dRAR⁺11] in small systems. The smooth entropies have also been used to investigate thermalization [HW11].

Finally, it has been shown [DR09] that the smooth entropy framework entails the information spectrum method by Han and Verdú [HV93, Han02] and its quantum generalization due to Hayashi, Nagaoka, and Ogawa [HN02, NH07, NO00].

9.2 Outlook and Open Questions

Some of the technical results in this thesis are new and applications remain unexplored. The entropic asymptotic equipartition property presented here has been improved from earlier work [TCR09] and now provides a converse bound for finite smoothing. This may help in the quest to prove strong converse statements for various tasks in quantum information theory, including channel capacities. These arguments apply to classical theory as well and it remains to be seen how these converse bounds compare to the literature.

In Chapter 7, Eq. (7.15), it was conjectured that the effective overlap is always smaller than the overlap. It remains an open question to show that this is true for general POVMs.

The entropic uncertainty relation has been generalized from previous work [TR11]. The lower bound on the uncertainty is now expressed using an effective overlap. In turn, this effective overlap can be bounded in terms of the maximal CHSH value that can be reached — using the same measurement setup — with an arbitrary second party. The CHSH value is measure of the non-locality of classical correlations produced by two parties. We thus believe that the generalized uncertainty relation provides a new avenue into device-independent quantum cryptography [HT11, LPT⁺11]. More generally, the applications of the uncertainty relation for smooth entropies in quantum cryptography are not fully explored yet. Another interesting extension is the bipartite uncertainty relation. It has potential applications in two-party quantum cryptography, which are unexplored.

Finally, the recent addition of a complete set of chain rules for the smooth entropies provides an important missing link and will lead to an even larger range of applicability of the framework.

Bibliography

- [AF04] R. Alicki and M. Fannes. Continuity of Quantum Conditional Information. *J. Phys. A: Math. Gen.*, 37(5): L55–L57, 2004. DOI: [10.1088/0305-4470/37/5/L01](https://doi.org/10.1088/0305-4470/37/5/L01).
- [AMAV07] K. M. R. Audenaert, L. Masanes, A. Acín, and F. Verstraete. Discriminating States: The Quantum Chernoff Bound. *Phys. Rev. Lett.*, 98(16), 2007. DOI: [10.1103/PhysRevLett.98.160501](https://doi.org/10.1103/PhysRevLett.98.160501).
- [ANSV08] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete. Asymptotic Error Rates in Quantum Hypothesis Testing. *Comm. Math. Phys.*, 279(1): 251–283, 2008. DOI: [10.1007/s00220-008-0417-5](https://doi.org/10.1007/s00220-008-0417-5).
- [BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. IEEE Int. Conf. on Comp., Sys. and Signal Process.*, pages 175–179, Bangalore, 1984. IEEE.
- [BBCM95] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer. Generalized Privacy Amplification. *IEEE Trans. on Inf. Theory*, 41(6): 1915–1923, 1995. DOI: [10.1109/18.476316](https://doi.org/10.1109/18.476316).
- [BBM92] C. H. Bennett, G. Brassard, and N. D. Mermin. Quantum cryptography without Bell’s theorem. *Phys. Rev. Lett.*, 68(5): 557–559, 1992. DOI: [10.1103/PhysRevLett.68.557](https://doi.org/10.1103/PhysRevLett.68.557).
- [BBR88] C. H. Bennett, G. Brassard, and J.-M. Robert. Privacy Amplification by Public Discussion. *SIAM J. on Comput.*, 17(2): 210, 1988. DOI: [10.1137/0217014](https://doi.org/10.1137/0217014).
- [BCC⁺10] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner. The Uncertainty Principle in the Presence of Quantum Memory. *Nat. Phys.*, 6(9): 659–662, 2010. DOI: [10.1038/nphys1734](https://doi.org/10.1038/nphys1734).

BIBLIOGRAPHY

- [BCR11] M. Berta, M. Christandl, and R. Renner. The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory. *Commun. Math. Phys.*, 306(3): 579–615, 2011.
DOI: [10.1007/s00220-011-1309-7](https://doi.org/10.1007/s00220-011-1309-7).
- [BD10] F. Buscemi and N. Datta. The Quantum Capacity of Channels With Arbitrarily Correlated Noise. *IEEE Trans. on Inf. Theory*, 56(3): 1447–1460, 2010.
DOI: [10.1109/TIT.2009.2039166](https://doi.org/10.1109/TIT.2009.2039166).
- [Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1: 195–200, 1964.
- [Ber08] M. Berta. Single-Shot Quantum State Merging. Master’s thesis, ETH Zurich, 2008. arXiv: [0912.4495](https://arxiv.org/abs/0912.4495).
- [BFS11] M. Berta, F. Furrer, and V. B. Scholz. The Smooth Entropy Formalism on von Neumann Algebras. 2011.
arXiv: [1107.5460](https://arxiv.org/abs/1107.5460).
- [BFW11] M. Berta, O. Fawzi, and S. Wehner. Quantum to Classical Randomness Extractors. 2011. arXiv: [1111.2026](https://arxiv.org/abs/1111.2026).
- [Bha97] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer, 1997.
- [Bha07] R. Bhatia. *Positive Definite Matrices*. Princeton Series in Applied Mathematics, 2007.
- [Bol72] L. Boltzmann. Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. In *Sitzungsberichte der Akademie der Wissenschaften zu Wien*, volume 66, pages 275–370, 1872.
- [BR11] N. J. Beaudry and R. Renner. An Intuitive Proof of the Data Processing Inequality. 2011. arXiv: [1107.0740](https://arxiv.org/abs/1107.0740).
- [Bur69] D. Bures. An extension of Kakutani’s theorem on infinite product measures to the tensor product of semifinite ω^* -algebras. *Trans. Amer. Math. Soc.*, 135: 199–212, 1969.
DOI: [10.1090/S0002-9947-1969-0236719-2](https://doi.org/10.1090/S0002-9947-1969-0236719-2).
- [Can01] R. Canetti. Universally Composable Security: A New Paradigm for Cryptographic Protocols. In *Proc. IEEE Int. Conf. on Cluster Comput.*, pages 136–145. IEEE, 2001.
DOI: [10.1109/SFCS.2001.959888](https://doi.org/10.1109/SFCS.2001.959888).

BIBLIOGRAPHY

- [CBKG02] N. Cerf, M. Bourennane, A. Karlsson, and N. Gisin. Security of Quantum Key Distribution Using d-Level Systems. *Phys. Rev. Lett.*, 88(12), 2002. DOI: [10.1103/PhysRevLett.88.127902](https://doi.org/10.1103/PhysRevLett.88.127902).
- [CHSH69] J. Clauser, M. Horne, A. Shimony, and R. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Phys. Rev. Lett.*, 23(15): 880–884, 1969. DOI: [10.1103/PhysRevLett.23.880](https://doi.org/10.1103/PhysRevLett.23.880).
- [CR08] R. Colbeck and R. Renner. Hidden Variable Models for Quantum Theory cannot have any Local Part. *Phys. Rev. Lett.*, 101(5), 2008. DOI: [10.1103/PhysRevLett.101.050403](https://doi.org/10.1103/PhysRevLett.101.050403).
- [CR10] R. Colbeck and R. Renner. No extension of quantum theory can have improved predictive power. *Nat. Commun.* 2:411, 2011. DOI: [10.1038/ncomms1416](https://doi.org/10.1038/ncomms1416).
- [Csi72] I. Csiszár. A class of measures of informativity of observation channels. *Periodica Math. Hungarica*, 2(1-4): 191–213, 1972. DOI: [10.1007/BF02018661](https://doi.org/10.1007/BF02018661).
- [CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991. DOI: [10.1002/047174882X](https://doi.org/10.1002/047174882X).
- [CW79] J. L. Carter and M. N. Wegman. Universal Classes of Hash Functions. *J. Comp. Syst. Sci.*, 18(2): 143–154, 1979. DOI: [10.1016/0022-0000\(79\)90044-8](https://doi.org/10.1016/0022-0000(79)90044-8).
- [CW05] M. Christandl and A. Winter. Uncertainty, Monogamy, and Locking of Quantum Correlations. *IEEE Trans. on Inf. Theory*, 51(9): 3159–3165, 2005. DOI: [10.1109/TIT.2005.853338](https://doi.org/10.1109/TIT.2005.853338).
- [CYGG11] P. Coles, L. Yu, V. Gheorghiu, and R. Griffiths. Information-theoretic treatment of tripartite systems and quantum channels. *Phys. Rev. A*, 83(6), 2011. DOI: [10.1103/PhysRevA.83.062338](https://doi.org/10.1103/PhysRevA.83.062338).
- [CYZ11] P. J. Coles, L. Yu, and M. Zwolak. Relative Entropy Derivation of the Uncertainty Principle with Quantum Side Information. 2011. [arXiv: 1105.4865](https://arxiv.org/abs/1105.4865).
- [Dat09] N. Datta. Min- and Max- Relative Entropies and a New Entanglement Monotone. *IEEE Trans. on Inf. Theory*, 55(6): 2816–2826, 2009. DOI: [10.1109/TIT.2009.2018325](https://doi.org/10.1109/TIT.2009.2018325).

BIBLIOGRAPHY

- [DBWR10] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner. The Decoupling Theorem. 2010. [arXiv: 1012.6044](#).
- [DD10] S. P. Desrosiers and F. Dupuis. Quantum Entropic Security and Approximate Quantum Encryption. *IEEE Trans. on Inf. Theory*, 56(7): 3455–3464, 2010. DOI: [10.1109/TIT.2010.2048488](#).
- [Deu83] D. Deutsch. Uncertainty in Quantum Measurements. *Phys. Rev. Lett.*, 50(9): 631–633, 1983. DOI: [10.1103/PhysRevLett.50.631](#).
- [DH11] N. Datta and M.-H. Hsieh. One-Shot Entanglement-Assisted Quantum and Classical Communication. 2011. [arXiv: 1105.3321](#).
- [DHB11] N. Datta, M.-H. Hsieh, and F. Brandao. Strong Converse Rates and an Example of Violation of the Strong Converse Property. 2011. [arXiv: 1106.3089](#).
- [DPVR09] A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan’s Extractor in the Presence of Quantum Side Information. 2009. [arXiv: 0912.5514](#).
- [DR09] N. Datta and R. Renner. Smooth Entropies and the Quantum Information Spectrum. *IEEE Trans. on Inf. Theory*, 55(6): 2807–2815, 2009. DOI: [10.1109/TIT.2009.2018340](#).
- [dRAR⁺11] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral. The Thermodynamic Meaning of Negative Entropy. *Nature*, 474(7349): 61–3, 2011. DOI: [10.1038/nature10123](#).
- [DRRV11] O. Dahlsten, R. Renner, E. Rieper, and V. Vedral. Inadequacy of von Neumann Entropy for Characterizing Extractable Work. *New J. Phys.*, 13(5): 053015, 2011. DOI: [10.1088/1367-2630/13/5/053015](#).
- [Dup09] F. Dupuis. *The Decoupling Approach to Quantum Information Theory*. PhD thesis, Université de Montréal, 2009. [arXiv: 1004.1641](#).
- [DW03] I. Devetak and A. Winter. Classical Data Compression with Quantum Side Information. *Phys. Rev. A*, 68(4), 2003. DOI: [10.1103/PhysRevA.68.042301](#).
- [Eke91] A. K. Ekert. Quantum cryptography based on Bell’s theorem. *Phys. Rev. Lett.*, 67(6): 661–663, 1991. DOI: [10.1103/PhysRevLett.67.661](#).

BIBLIOGRAPHY

- [EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47: 777–780, 1935.
- [FAbR11] F. Furrer, J. Åberg, and R. Renner. Min- and Max-Entropy in Infinite Dimensions. *Commun. Math. Phys.*, 306(1): 165–186, 2011. DOI: [10.1007/s00220-011-1282-1](https://doi.org/10.1007/s00220-011-1282-1).
- [Fan73] M. Fannes. A Continuity Property of the Entropy Density for Spin Lattice Systems, 1973. DOI: [10.1007/BF01646490](https://doi.org/10.1007/BF01646490).
- [FS08] S. Fehr and C. Schaffner. Randomness Extraction Via Delta-Biased Masking in the Presence of a Quantum Attacker. In *Proc. TCC*, volume 4948 of *LNCS*, pages 465–481, New York, 2008. Springer. arXiv: [0706.2606v2](https://arxiv.org/abs/0706.2606v2).
- [Fuc96] C. A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. Phd thesis, University of New Mexico, 1996. arXiv: [quant-ph/9601020v1](https://arxiv.org/abs/quant-ph/9601020v1).
- [Gal79] R. G. Gallager. Source Coding with Side Information and Universal Coding. In *Proc. IEEE ISIT*, volume 21, Ronneby, Sweden,, 1979. IEEE.
- [Gib76] J. W. Gibbs. On the Equilibrium of Heterogeneous Substances. *Transactions of the Connecticut Academy*, III: 108–248, 1876.
- [GKK⁺07] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential Separation for one-way Quantum Communication Complexity, with Applications to Cryptography. In *Proc. ACM STOC*, pages 516–525. ACM Press, 2007. arXiv: [arXiv:quant-ph/0611209v3](https://arxiv.org/abs/quant-ph/0611209v3).
- [GLN04] A. Gilchrist, N. K. Langford, and M. A. Nielsen. Distance Measures to Compare Real and Ideal Quantum Processes, 2004. arXiv: [0408063](https://arxiv.org/abs/0408063).
- [Hal95] M. Hall. Information Exclusion Principle for Complementary Observables. *Phys. Rev. Lett.*, 74(17): 3307–3311, 1995. DOI: [10.1103/PhysRevLett.74.3307](https://doi.org/10.1103/PhysRevLett.74.3307).
- [Han02] T. S. Han. *Information-Spectrum Methods in Information Theory*. Springer, 2002.
- [Har28] R. V. L. Hartley. Transmission of Information. *Bell Syst. Tech. J.*, 7(3): 535–563, 1928.

BIBLIOGRAPHY

- [Har01] L. Hardy. Quantum Theory From Five Reasonable Axioms. page 34, 2001. [arXiv: quant-ph/0101012](#).
- [Hay06] M. Hayashi. *Quantum Information — An Introduction*. Springer, 2006.
- [Hei27] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Phys.*, 43(3-4): 172–198, 1927.
- [HHH96] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223(1-2): 1–8, 1996. DOI: [10.1016/S0375-9601\(96\)00706-2](#).
- [Hir57] I. I. Hirschman. A Note on Entropy. *Am. J. Math.*, 79(1): 152–156, 1957.
- [HK69] K. E. Hellwig and K. Kraus. Pure Operations and Measurements. *Commun. Math. Phys.*, 11(3): 214–220, 1969. DOI: [10.1007/BF01645807](#).
- [HK70] K. E. Hellwig and K. Kraus. Operations and Measurements II. *Commun. Math. Phys.*, 16(2): 142–147, 1970. DOI: [10.1007/BF01646620](#).
- [HMPB11] F. Hiai, M. Mosonyi, D. Petz, and C. Bény. Quantum f-Divergences and Error Correction. *Rev. Math. Phys.*, 23(07): 691–747, 2011. DOI: [10.1142/S0129055X11004412](#).
- [HN02] M. Hayashi and H. Nagaoka. General Formulas for Capacity of Classical-Quantum Channels. *IEEE Trans. on Inf. Theory*, 49: 1753–1768, 2002. DOI: [10.1109/TIT.2003.813556](#).
- [Hol73] A. S. Holevo. Some Estimates for Information Quantity Transmitted by Quantum Communication Channel. *Probl. Inform. Transm.*, 9(3): 3–11, 1973.
- [HOW05] M. Horodecki, J. Oppenheim, and A. Winter. Partial Quantum Information. *Nature*, 436(7051): 673–6, 2005. DOI: [10.1038/nature03909](#).
- [HOW06] M. Horodecki, J. Oppenheim, and A. Winter. Quantum State Merging and Negative Information. *Commun. Math. Phys.*, 269(1): 107–136, 2006. DOI: [10.1007/s00220-006-0118-x](#).
- [HP03] F. Hansen and G. K. Pedersen. Jensen’s Operator Inequality. *B. Lond. Math. Soc.*, 35(4): 553–564, 2003.

BIBLIOGRAPHY

- [HR11] T. Holenstein and R. Renner. On the Randomness of Independent Experiments, 2011. DOI: [10.1109/TIT.2011.2110230](https://doi.org/10.1109/TIT.2011.2110230).
- [HT11] E. Hänggi and M. Tomamichel. The Link between Uncertainty Relations and Non-Locality. 2011. arXiv: [1108.5349](https://arxiv.org/abs/1108.5349).
- [HV93] T. S. Han and S. Verdú. Approximation Theory of Output Statistics. In *Proc. ISIT*, pages 153–153. IEEE, 1993. DOI: [10.1109/ISIT.1993.748468](https://doi.org/10.1109/ISIT.1993.748468).
- [HW11] F. Hutter and S. Wehner. When does a Quantum Mechanical System depend on the Initial Conditions of the System or the Environment? arXiv: [1111.3080](https://arxiv.org/abs/1111.3080)
- [ILL89] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *Proc. ACM STOC*, pages 12–24. ACM Press, 1989. DOI: [10.1145/73007.73009](https://doi.org/10.1145/73007.73009).
- [IZ89] R. Impagliazzo and D. Zuckerman. How to Recycle Random Bits. In *Proc. IEEE Symp. on Found. of Comp. Sc.*, pages 248–253, 1989. DOI: [10.1109/SFCS.1989.63486](https://doi.org/10.1109/SFCS.1989.63486).
- [Jam72] A. Jamiolkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3(4): 275–278, 1972.
- [Kle31] O. Klein. Zur Quantenmechanischen Begründung des zweiten Hauptsatzes der Wärmelehre. *Z. Phys.*, 72(11-12): 767–775, 1931. DOI: [10.1007/BF01341997](https://doi.org/10.1007/BF01341997).
- [KMR05] R. König, U. Maurer, and R. Renner. On the Power of Quantum Memory. *IEEE Trans. on Inf. Theory*, 51(7): 2391–2401, 2005. DOI: [10.1109/TIT.2005.850087](https://doi.org/10.1109/TIT.2005.850087).
- [KP02] M. Krishna and K. R. Parthasarathy. An Entropic Uncertainty Principle for Quantum Measurements. *Indian J. Stat.*, 64(3): 842–851, 2002.
- [KR11] R. König and R. Renner. Sampling of Min-Entropy Relative to Quantum Knowledge. *IEEE Trans. on Inf. Theory*, 57(7): 4760–4787, 2011. DOI: [10.1109/TIT.2011.2146730](https://doi.org/10.1109/TIT.2011.2146730).
- [KRBM07] R. König, R. Renner, A. Bariska, and U. Maurer. Small Accessible Quantum Information Does Not Imply Security. *Phys. Rev. Lett.*, 98(14), 2007. DOI: [10.1103/PhysRevLett.98.140502](https://doi.org/10.1103/PhysRevLett.98.140502).

BIBLIOGRAPHY

- [KRS09] R. König, R. Renner, and C. Schaffner. The Operational Meaning of Min- and Max-Entropy. *IEEE Trans. on Inf. Theory*, 55(9): 4337–4347, 2009. DOI: [10.1109/TIT.2009.2025545](https://doi.org/10.1109/TIT.2009.2025545).
- [KS68] S. Kochen and E. P. Specker. The Problem of Hidden Variables in Quantum Mechanics. *Indiana Univ. Math. J.*, 17(1): 59–87, 1968.
- [KWW09] R. Koenig, S. Wehner, and J. Wullschleger. Unconditional Security from Noisy Quantum Storage. page 25, 2009. arXiv: [0906.1030](https://arxiv.org/abs/0906.1030).
- [Löf04] J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *Proc. of the CACSD Conf.*, Taipei, Taiwan, 2004. Online: <http://users.isy.liu.se/johanl/yalmip>.
- [LPT⁺11] C. C. W. Lim, C. Portmann, M. Tomamichel, N. Gisin, and R. Renner. Unpublished Manuscript. 2011.
- [LSS11] T. P. Le, L. Sheridan, and V. Scarani. Tomographic Quantum Cryptography Protocols are Reference Frame Independent. 2011. arXiv: [1109.2510](https://arxiv.org/abs/1109.2510).
- [McI87] J. McInnes. Cryptography Using Weak Sources of Randomness. *Tech. Report, U. of Toronto*, 1987.
- [MD09] M. Mosonyi and N. Datta. Generalized Relative Entropies and the Capacity of Classical-Quantum Channels. *J. Math. Phys.*, 50(7): 072104, 2009. DOI: [10.1063/1.3167288](https://doi.org/10.1063/1.3167288).
- [MU88] H. Maassen and J. Uffink. Generalized entropic uncertainty relations. *Phys. Rev. Lett.*, 60(12): 1103–1106, 1988. DOI: [10.1103/PhysRevLett.60.1103](https://doi.org/10.1103/PhysRevLett.60.1103).
- [NC00] M. A. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [NH07] H. Nagaoka and M. Hayashi. An Information-Spectrum Approach to Classical and Quantum Hypothesis Testing for Simple Hypotheses. *IEEE Trans. on Inf. Theory*, 53(2): 534–549, 2007. DOI: [10.1109/TIT.2006.889463](https://doi.org/10.1109/TIT.2006.889463).
- [NO00] H. Nagaoka and T. Ogawa. Strong Converse and Stein’s Lemma in Quantum Hypothesis Testing. *IEEE Trans. on Inf. Theory*, 46(7): 2428–2433, 2000. DOI: [10.1109/18.887855](https://doi.org/10.1109/18.887855).

BIBLIOGRAPHY

- [NP05] M. A. Nielsen and D. Petz. A Simple Proof of the Strong Subadditivity Inequality. *Quant. Inf. Comput.*, 5(6): 507–513, 2005. [arXiv: quant-ph/0408130](#).
- [Nyq28] H. Nyquist. Certain Topics in Telegraph Transmission Theory. *Trans. AIEE*, 47(2): 617–644, 1928.
DOI: [10.1109/T-AIEE.1928.5055024](#).
- [Ooh94] Y. Oohama. Universal Coding for the Slepian-Wolf Data Compression system and the strong converse theorem. *IEEE Trans. on Inf. Theory*, 40(6): 1908–1919, 1994.
DOI: [10.1109/18.340465](#).
- [OP93] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer, 1993.
- [Per96] A. Peres. Separability Criterion for Density Matrices. *Phys. Rev. Lett.*, 77(8): 1413–1415, 1996.
DOI: [10.1103/PhysRevLett.77.1413](#).
- [Pet84] D. Petz. Quasi-Entropies for Finite Quantum Systems. *Rep. Math. Phys.*, 23: 57–65, 1984.
- [Pet10] D. Petz. From f-Divergence to Quantum Quasi-Entropies and Their Use. *Entropy*, 12(3): 304–325, 2010.
DOI: [10.3390/e12030304](#).
- [Pol10] Y. Polyanskiy. *Channel Coding: Non-Asymptotic Fundamental Limits*. PhD thesis, Princeton University, 2010.
- [PPV10] Y. Polyanskiy, H. V. Poor, and S. Verdú. Channel Coding Rate in the Finite Blocklength Regime. *IEEE Trans. on Inf. Theory*, 56(5): 2307–2359, 2010.
DOI: [10.1109/TIT.2010.2043769](#).
- [PW01] B. Pfitzmann and M. Waidner. A Model for Asynchronous Reactive Systems and its Application to Secure Message Transmission. In *Proc. IEEE Symp. on Security and Privacy*, pages 184–200. IEEE Comput. Soc, 2001.
DOI: [10.1109/SECPRI.2001.924298](#).
- [Rén61] A. Rényi. On Measures of Information and Entropy. In *Proc. Symp. on Math., Stat. and Probability*, pages 547–561, Berkeley, 1961. University of California Press.
- [Ras06] A. E. Rastegin. Sine Distance for Quantum States, 2006.
[arXiv: quant-ph/0602112](#).

BIBLIOGRAPHY

- [RB09] J. Renes and J.-C. Boileau. Conjectured Strong Complementary Information Tradeoff. *Phys. Rev. Lett.*, 103(2), 2009. DOI: [10.1103/PhysRevLett.103.020402](https://doi.org/10.1103/PhysRevLett.103.020402).
- [Ren05] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005. arXiv: [quant-ph/0512258](https://arxiv.org/abs/quant-ph/0512258).
- [RK05] R. Renner and R. König. Universally Composable Privacy Amplification Against Quantum Adversaries. In *Proc. TCC*, volume 3378 of *LNCS*, pages 407–425, Cambridge, USA, 2005. DOI: [10.1007/978-3-540-30576-7_22](https://doi.org/10.1007/978-3-540-30576-7_22).
- [Rob29] H. P. Robertson. The Uncertainty Principle. *Phys. Rev.*, 34(1): 163–164, 1929. DOI: [10.1103/PhysRev.34.163](https://doi.org/10.1103/PhysRev.34.163).
- [RR10] J. M. Renes and R. Renner. One-Shot Classical Data Compression with Quantum Side Information and the Distillation of Common Randomness or Secret Keys. 2010. arXiv: [1008.0452](https://arxiv.org/abs/1008.0452).
- [RW04] R. Renner and S. Wolf. Smooth Rényi Entropy and Applications. In *Proc. ISIT*, pages 232–232. IEEE, 2004. DOI: [10.1109/ISIT.2004.1365269](https://doi.org/10.1109/ISIT.2004.1365269).
- [RW05] R. Renner and S. Wolf. Simple and Tight Bounds for Information Reconciliation and Privacy Amplification. In *Proc. ASIACRYPT*, volume 3788 of *LNCS*, pages 199–216. Springer, 2005.
- [Sch95] B. Schumacher. Quantum Coding. *Phys. Rev. A*, 51(4): 2738–2747, 1995. DOI: [10.1103/PhysRevA.51.2738](https://doi.org/10.1103/PhysRevA.51.2738).
- [SDTR11] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner. Decoupling with Unitary Almost Two-Designs. 2011. arXiv: [1109.4348](https://arxiv.org/abs/1109.4348).
- [SED] SeDuMi. Online: <http://sedumi.ie.lehigh.edu/>.
- [Sha48] C. Shannon. A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, 27: 379–423, 1948.
- [Sha51] C. Shannon. Prediction and Entropy of Printed English. *Bell Syst. Tech. J.*, 30: 50 – 64, 1951.
- [SR08] V. Scarani and R. Renner. Security Bounds for Quantum Cryptography with Finite Resources. In *Proc. TQC*, volume 5106 of *LNCS*, pages 83–95. Springer, 2008. arXiv: [0806.0120](https://arxiv.org/abs/0806.0120).

BIBLIOGRAPHY

- [Sti55] W. F. Stinespring. Positive Functions On C*-Algebras. *Proc. Am. Math. Soc.*, 6: 211–216, 1955.
DOI: [10.1090/S0002-9939-1955-0069403-4](https://doi.org/10.1090/S0002-9939-1955-0069403-4).
- [SW73] D. Slepian and J. Wolf. Noiseless Coding of Correlated Information Sources. *IEEE Trans. on Inf. Theory*, 19(4): 471–480, 1973. DOI: [10.1109/TIT.1973.1055037](https://doi.org/10.1109/TIT.1973.1055037).
- [SW08] V. B. Scholz and R. F. Werner. Tsirelson’s Problem. 2008.
arXiv: [0812.4305](https://arxiv.org/abs/0812.4305).
- [TCR09] M. Tomamichel, R. Colbeck, and R. Renner. A Fully Quantum Asymptotic Equipartition Property. *IEEE Trans. on Inf. Theory*, 55(12): 5840–5847, 2009.
DOI: [10.1109/TIT.2009.2032797](https://doi.org/10.1109/TIT.2009.2032797).
- [TCR10] M. Tomamichel, R. Colbeck, and R. Renner. Duality Between Smooth Min- and Max-Entropies. *IEEE Trans. on Inf. Theory*, 56(9): 4674–4681, 2010.
DOI: [10.1109/TIT.2010.2054130](https://doi.org/10.1109/TIT.2010.2054130).
- [TLGR11] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner. Tight Finite-Key Analysis for Quantum Cryptography. *Nat. Commun.*, 3(634), 2012. DOI: [10.1038/ncomms1631](https://doi.org/10.1038/ncomms1631).
- [TR11] M. Tomamichel and R. Renner. Uncertainty Relation for Smooth Entropies. *Phys. Rev. Lett.*, 106(11), 2011.
DOI: [10.1103/PhysRevLett.106.110506](https://doi.org/10.1103/PhysRevLett.106.110506).
- [TSSR11] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover Hashing Against Quantum Side Information. *IEEE Trans. on Inf. Theory*, 57(8): 5524–5535, 2011.
DOI: [10.1109/TIT.2011.2158473](https://doi.org/10.1109/TIT.2011.2158473).
- [Uhl85] A. Uhlmann. The Transition Probability for States of Star-Algebras. *Ann. Phys.*, 497(4): 524–532, 1985.
- [Unr10] D. Unruh. Universally Composable Quantum Multi-party Computation. In *Proc. EUROCRYPT*, volume 6110 of *LNCS*, pages 486–505. Springer, 2010.
DOI: [10.1007/978-3-642-13190-5](https://doi.org/10.1007/978-3-642-13190-5).
- [VDTR11] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner. Unpublished Manuscript. 2011.
- [Vit11] A. Vitanov. *Smooth Min- And Max-Entropy Calculus: Chain Rules*. Master’s thesis, ETH Zurich, 2011.

BIBLIOGRAPHY

- [vN32] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, 1932.
- [Wat08] J. Watrous. Theory of Quantum Information, Lecture Notes, 2008. Online: <http://www.cs.uwaterloo.ca/~watrous/>.
- [WCR09] L. Wang, R. Colbeck, and R. Renner. Simple Channel Coding Bounds. In *Proc. IEEE ISIT*, pages 1804–1808. IEEE, 2009. DOI: [10.1109/ISIT.2009.5205312](https://doi.org/10.1109/ISIT.2009.5205312).
- [WTHR11] S. Winkler, M. Tomamichel, S. Hengl, and R. Renner. Impossibility of Growing Quantum Bit Commitments. *Phys. Rev. Lett.*, 107(9), 2011. DOI: [10.1103/PhysRevLett.107.090502](https://doi.org/10.1103/PhysRevLett.107.090502).
- [Win99] A. Winter. *Coding Theorems of Quantum Information Theory*. PhD thesis, Universität Bielefeld, 1999. arXiv: [quant-ph/9907077](https://arxiv.org/abs/quant-ph/9907077).
- [Wul08] J. Wullschleger. Private Communication. 2008.
- [WW08] S. Wehner and J. Wullschleger. Composable Security in the Bounded-Quantum-Storage Model. In *Proc. ICALP*, volume 5126 of *LNCS*, pages 604–615. Springer, 2008. DOI: [10.1007/978-3-540-70583-3](https://doi.org/10.1007/978-3-540-70583-3).
- [WW10] S. Wehner and A. Winter. Entropic Uncertainty Relations—A Survey. *New J. Phys.*, 12(2): 025009, 2010. DOI: [10.1088/1367-2630/12/2/025009](https://doi.org/10.1088/1367-2630/12/2/025009).
- [WZ82] W. K. Wootters and W. H. Zurek. A Single Quantum Cannot be Cloned. *Nature*, 299(5886): 802–803, 1982. DOI: [10.1038/299802a0](https://doi.org/10.1038/299802a0).

Appendix A

Various Lemmas

A.1 Two Lemmas for Tensor Spaces

The following Lemma gives an operator inequality that relates the *marginal* states on one Hilbert space before and after a *CPM* is applied on the other space.

Lemma A.1. *Let $M \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H}')$ and let \mathcal{E} be a CPM from \mathcal{H}' to \mathcal{H}'' . Then,*

$$\mathrm{tr}_{\mathcal{H}''}(\mathcal{E}[M]) \leq \|\mathcal{E}^\dagger[\mathbb{1}]\|_\infty \mathrm{tr}_{\mathcal{H}'}(M).$$

Proof. We write \mathcal{E} in its *Kraus representation*, i.e. $\mathcal{E}[M] = \sum_k E_k M E_k^\dagger$, where $E_k \in \mathcal{L}(\mathcal{H}', \mathcal{H}'')$. Due to the linearity and cyclicity of the partial trace, we have

$$\mathrm{tr}_{\mathcal{H}''}(\mathcal{E}[M]) = \mathrm{tr}_{\mathcal{H}'}\left(\sum_k E_k^\dagger E_k M\right) = \mathrm{tr}_{\mathcal{H}'}(\mathcal{E}^\dagger[\mathbb{1}] M)$$

We introduce the operator $R = \mathbb{1}\|\mathcal{E}^\dagger[\mathbb{1}]\|_\infty - \mathcal{E}^\dagger[\mathbb{1}] \geq 0$ and note that $\mathrm{tr}_{\mathcal{H}'}(\sqrt{R}M\sqrt{R}) \geq 0$ and, thus,

$$\mathrm{tr}_{\mathcal{H}'}(\mathcal{E}^\dagger[\mathbb{1}] M) \leq \mathrm{tr}_{\mathcal{H}'}((\mathcal{E}^\dagger[\mathbb{1}] + R)M) = \|\mathcal{E}^\dagger[\mathbb{1}]\|_\infty \mathrm{tr}_{\mathcal{H}'}(M). \quad \square$$

In particular, this lemma implies that *TP-CPMs* do not affect the marginal state on another space. To see this, note that the bound evaluates to $\mathrm{tr}_{\mathcal{H}''}(\mathcal{E}[M]) \leq \mathrm{tr}_{\mathcal{H}'}(M)$. Hence, the operator $\mathrm{tr}_{\mathcal{H}'}(M) - \mathrm{tr}_{\mathcal{H}''}(\mathcal{E}[M])$ is positive and has vanishing trace, implying that the two marginal states are in fact equal. Moreover, if the map \mathcal{E} is trace non-increasing, we still find $\mathrm{tr}_{\mathcal{H}''}(\mathcal{E}[M]) \leq \mathrm{tr}_{\mathcal{H}'}(M)$ as its adjoint is sub-unital.

As a further example, consider the CPM $\mathcal{L} : M \mapsto LML^\dagger$ that conjugates M with an arbitrary linear operator $L \in \mathcal{L}(\mathcal{H}', \mathcal{H}'')$. In this case, $\mathcal{L}^\dagger[\mathbb{1}]$ evaluates to $L^\dagger L = |L|^2$ and we find

$$\mathrm{tr}_{\mathcal{H}''}(\mathcal{L}[M]) \leq \| |L|^2 \|_\infty \mathrm{tr}_{\mathcal{H}'}(M) = \|L\|_\infty^2 \mathrm{tr}_{\mathcal{H}'}(M). \quad (\text{A.1})$$

A. VARIOUS LEMMAS

To simplify the expression, we used that $\|N^2\|_\infty = \|N\|_\infty^2$ for $N \in \mathcal{P}(H)$ and that $\|L\| = \|L\|$ due to the *unitary invariance* of the norm applied to the *polar decomposition* of L .

Moreover, the following operator inequality holds.

Lemma A.2. *Let $M \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H}')$. Then, $M \leq \dim\{\mathcal{H}'\} \operatorname{tr}_{\mathcal{H}'}(M) \otimes \mathbb{1}_{\mathcal{H}'}$.*

Proof. Since M has an eigenvalue decomposition with positive eigenvalues, it is sufficient to prove the property for an operator M of the form $M = |\varphi\rangle\langle\varphi|$, where $\varphi \in \mathcal{H} \otimes \mathcal{H}'$. The general result then follows by linearity.

Let $X = \operatorname{tr}_{\mathcal{H}'}(M)$ and $\Lambda = X^{-1/2}|\varphi\rangle\langle\varphi|X^{-1/2}$, where we take the *generalized inverse* of X . Since $\operatorname{tr}(\Lambda) = \operatorname{rank}\{X\} \leq \dim\{\mathcal{H}'\}$ due to the Schmidt decomposition, we have $\Lambda \leq \dim\{\mathcal{H}'\} \mathbb{1}$. The statement of the lemma then follows by conjugating this inequality with $X^{1/2}$. \square

A.2 Entropies of Coherent Classical States

The following two technical lemmas are useful when *coherent classical* states are discussed.

Lemma A.3. *Let $\tau \in \mathcal{S}_\leq(\mathcal{H}_{XX'AB})$. Then, the corresponding coherent classical state $\rho = \Pi_{XX'}\tau\Pi_{XX'}$ satisfies*

$$H_{\min}(XA|X'B)_\tau \leq H_{\min}(XA|X'B)_\rho \leq H_{\min}(A|XX'B)_\rho.$$

Proof. By the definition of the min-entropy, there exists a state $\sigma \in \mathcal{S}_=(\mathcal{H}_{X'B})$ such that

$$\tau_{XX'AB} \leq 2^{-H_{\min}(XA|X'B)_\tau} \mathbb{1}_{XA} \otimes \sigma_{X'B}.$$

Therefore, by conjugation with $\Pi_{XX'}$, we have

$$\begin{aligned} \rho_{XX'AB} &\leq 2^{-H_{\min}(XA|X'B)_\tau} \mathbb{1}_A \otimes \sum_x |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \langle x|\sigma_{X'B}|x\rangle_{X'} \\ &\leq 2^{-H_{\min}(XA|X'B)_\tau} \mathbb{1}_{AX} \otimes \mathcal{M}_{X'}[\sigma_{X'B}]. \end{aligned}$$

Since the measurement $\mathcal{M}_{X'}$ is trace-preserving, $\mathcal{M}_{X'}[\sigma]$ is a candidate for the optimization of $H_{\min}(XA|X'B)_\rho$ and the first inequality of the lemma follows. The second inequality follows by a similar argument, where $\sigma_{X'B}$ is chosen to maximize $H_{\min}(XA|X'B)_\rho$ and the state

$$\sigma_{XX'B} = \sum_x |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \langle x|\sigma_{X'B}|x\rangle_{X'}$$

is a candidate for the optimization of $H_{\min}(A|XX'B)$. \square

A.3 Selected Relations between Entropies

Lemma A.4. *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{XX'AB})$ be coherent classical between X and X' . Then,*

$$H_{\max}(XA|X'B)_{\rho} \leq H_{\max}(A|XX'B)_{\rho}.$$

Proof. Since the coherent classical state ρ commutes with $\Pi_{XX'}$, there exists a state $\sigma \in \mathcal{S}_{=}(\mathcal{H}_{X'B})$ such that

$$\begin{aligned} 2^{H_{\max}(XA|X'B)} &= \left\| \sqrt{\rho_{XX'AB}} \sqrt{\mathbb{1}_{XA} \otimes \sigma_{X'B}} \right\|_1 \\ &= \left\| \sqrt{\rho_{XX'AB}} \sqrt{\mathbb{1}_A \otimes \Pi_{XX'}(\mathbb{1}_X \otimes \sigma_{X'B}) \Pi_{XX'}} \right\|_1 \\ &\leq 2^{H_{\max}(A|XX'B)}. \end{aligned}$$

The last inequality follows from the fact that $\text{tr}(\Pi_{XX'}(\mathbb{1}_X \otimes \sigma_{X'B}) \Pi_{XX'}) = \text{tr}(\mathcal{M}_{X'}[\sigma]) = 1$, which makes this state a candidate for the optimization in $H_{\max}(A|XX'B)$. \square

A.3 Selected Relations between Entropies

The first lemma appeared in [TSSR11] and relates the *min-entropy* and the relative entropy where the conditioning is done on ρ .

Lemma A.5. *Let $\varepsilon > 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ be pure. Then, there exists a projector Π_{AC} on \mathcal{H}_{AC} and a state $\tilde{\rho} = \Pi \rho \Pi$ such that $\tilde{\rho} \in \mathcal{B}_p^{\varepsilon}(\rho)$ and*

$$S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B) \geq H_{\min}(A|B)_{\rho} - \log \frac{2}{\varepsilon^2}.$$

Proof. The proof is structured as follows: First, we give a lower bound on the entropy $S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B)$ in terms of $H_{\min}(A|B)_{\rho}$ and a projector Π_B that is a dual projector of Π_{AC} in the sense described below. We then find a lower bound on the purified distance between ρ and $\tilde{\rho}$ in terms of Π_B and define Π_B (and, thus, Π_{AC}) such that this distance does not exceed ε .

Let λ and σ be the pair that optimizes the min-entropy $H_{\min}(A|B)$, i.e. $H_{\min}(A|B)_{\rho} = S_{\min}(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) = \lambda$. We have $\tilde{\rho}_B \leq \rho_B$ by definition of $\tilde{\rho}$. Hence, $S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B)$ is finite and can be written as

$$2^{-S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B)} = \left\| \rho_B^{-1/2} \tilde{\rho}_{AB} \rho_B^{-1/2} \right\|_{\infty}. \quad (\text{A.2})$$

We bound this expression using the dual projector of Π_{AC} , Π_B , i.e. the projector that satisfies (cf. Lemma 2.1)

$$\Pi_B |\Gamma\rangle = \Pi_{AC} |\Gamma\rangle, \quad \text{where} \quad |\Gamma\rangle = \rho_B^{-1/2} |\rho\rangle = \rho_{AC}^{-1/2} |\rho\rangle = \text{vec}(\Pi^{\rho_B})$$

is the unnormalized fully entangled state between $\text{supp}\{\rho_B\}$ and $\text{supp}\{\rho_{AC}\}$. We also use that $\rho_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_B$ by definition of the min-entropy. Thus,

A. VARIOUS LEMMAS

we bound the rhs. of (A.2) as

$$\begin{aligned}
\text{rhs.} &= \left\| \text{tr}_C \left((\Pi_{AC} \otimes \rho_B^{-1/2}) \rho_{ABC} (\Pi_{AC} \otimes \rho_B^{-1/2}) \right) \right\|_\infty \\
&= \left\| \Pi_B \rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2} \Pi_B \right\|_\infty \\
&\leq 2^{-\lambda} \left\| \mathbb{1}_A \otimes (\Pi_B \rho_B^{-1/2} \sigma_B \rho_B^{-1/2} \Pi_B) \right\|_\infty \\
&= 2^{-\lambda} \left\| \Pi_B \Lambda_B \Pi_B \right\|_\infty,
\end{aligned}$$

where, in the last step, we introduced the Hermitian operator $\Lambda_B := \rho_B^{-1/2} \sigma_B \rho_B^{-1/2}$. Taking the logarithm on both sides leads to

$$S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B) \geq H_{\min}(A|B)_\rho - \log \left\| \Pi_B \Lambda_B \Pi_B \right\|_\infty. \quad (\text{A.3})$$

We use Lemma 3.8 to bound the distance between ρ_{ABC} and $\tilde{\rho}_{ABC}$, namely

$$P(\rho, \tilde{\rho}) \leq \sqrt{2 \text{tr}(\Pi_{AC}^\perp \rho_{ABC})} = \sqrt{2 \text{tr}(\Pi_B^\perp \rho_B)}, \quad (\text{A.4})$$

where the Π_{AC}^\perp and Π_B^\perp are the orthogonal complements of Π_{AC} and Π_B , respectively.

Clearly, the optimal choice of Π_B will cut off the largest eigenvalues of Λ in (A.3) while keeping ρ and $\tilde{\rho}$ close. We thus define Π_B to be the minimum rank projector onto the smallest eigenvalues of Λ such that $\text{tr}(\Pi_B \rho_B) \geq \text{tr}(\rho) - \varepsilon^2/2$ or, equivalently, $\text{tr}(\Pi_B^\perp \rho_B) \leq \varepsilon^2/2$. This definition immediately implies that $\rho \approx_\varepsilon \tilde{\rho}$ and it remains to find an upper bound on $\left\| \Pi_B \Lambda_B \Pi_B \right\|_\infty$.

Let Π'_B be the projector onto the largest remaining eigenvalue in the operator $\Pi_B \Lambda_B \Pi_B$ and note that Π'_B and Π_B^\perp commute with Λ_B . Then,

$$\left\| \Pi_B \Lambda_B \Pi_B \right\|_\infty = \text{tr}(\Pi'_B \Lambda_B) = \min_{\mu} \frac{\text{tr}(\mu_B (\Pi_B^\perp + \Pi'_B) \Lambda_B)}{\text{tr}(\mu)},$$

where μ is minimized over all positive operators in the support of $\Pi_B^\perp + \Pi'_B$. Fixing instead $\mu_B = (\Pi_B^\perp + \Pi'_B) \rho_B (\Pi_B^\perp + \Pi'_B)$, we find

$$\left\| \Pi_B \Lambda_B \Pi_B \right\|_\infty \leq \frac{\text{tr}(\Gamma_B^{1/2} \rho_B \Lambda_B^{1/2} (\Pi_B^\perp + \Pi'_B))}{\text{tr}((\Pi_B^\perp + \Pi'_B) \rho_B)} \leq \frac{\text{tr}(\Gamma_B^{1/2} \rho_B \Lambda_B^{1/2})}{\text{tr}((\Pi_B^\perp + \Pi'_B) \rho_B)} \leq \frac{2}{\varepsilon^2}.$$

In the last step we used that

$$\text{tr}(\Lambda_B^{1/2} \rho_B \Lambda_B^{1/2}) = \text{tr}(\sigma_B) = 1 \quad \text{and} \quad \text{tr}((\Pi_B^\perp + \Pi'_B) \rho_B) \geq \frac{\varepsilon^2}{2}$$

by definition of Π_B^\perp . This concludes the proof. \square

In addition, we need the following extension of this result to the smooth min-entropy.

A.3 Selected Relations between Entropies

Lemma A.6. *Let $\varepsilon > 0, \varepsilon' \geq 0$ and $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Then, there exists a state $\bar{\rho} \in \mathcal{B}^{\varepsilon+2\varepsilon'}(\rho)$ such that*

$$S_{\min}(\bar{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B) \geq H_{\min}^{\varepsilon'}(A|B)_{\rho} - \log \frac{2}{\varepsilon^2}.$$

Proof. Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ and $\hat{\rho} \in \mathcal{B}^{\varepsilon'}(\rho_{ABC})$ be pure states such that $H_{\min}^{\varepsilon'}(A|B)_{\rho} = H_{\min}(A|B)_{\hat{\rho}}$. We apply Lemma A.5 to this state to get

$$S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \hat{\rho}_B) \geq H_{\min}^{\varepsilon'}(A|B)_{\rho} - \log \frac{2}{\varepsilon^2}, \quad \text{where } |\tilde{\rho}\rangle = \Pi_{AC}|\hat{\rho}\rangle, \tilde{\rho} \approx_{\varepsilon} \hat{\rho}.$$

Using Lemma 3.7, we define the operator F_B with the property $F_B \hat{\rho}_B F_B^{\dagger} = \rho_B$; hence $F_B \hat{\rho}_{ABC} F_B^{\dagger} \approx_{\varepsilon'} \hat{\rho}_{ABC}$. Applying this to the defining operator inequality of the relative entropy above leads to

$$\tilde{\rho}_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \hat{\rho}_B \implies \bar{\rho}_{AB} := F_B \tilde{\rho}_{AB} F_B^{\dagger} \leq 2^{-\lambda} \mathbb{1}_A \otimes \rho_B$$

and, thus, $S_{\min}(\tilde{\rho}_{AB} \| \mathbb{1}_A \otimes \hat{\rho}_B) \leq S_{\min}(\bar{\rho}_{AB} \| \mathbb{1}_A \otimes \rho_B)$. Furthermore, $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ since $\text{tr}(\bar{\rho}) = \text{tr}(F_B \tilde{\rho}_B F_B^{\dagger}) \leq \text{tr}(F_B \hat{\rho}_B F_B^{\dagger}) = \text{tr}(\rho_B) \leq 1$. Hence, it remains to bound $P(\bar{\rho}, \rho) \leq P(\bar{\rho}, \tilde{\rho}) + P(\tilde{\rho}, \hat{\rho}) + P(\hat{\rho}, \rho) \leq P(\bar{\rho}, \tilde{\rho}) + \varepsilon + \varepsilon'$ and

$$\begin{aligned} P(\bar{\rho}, \tilde{\rho}) &= P((F_B \otimes \Pi_{AC}) \hat{\rho}_{ABC} (F_B \otimes \Pi_{AC}), \Pi_{AC} \hat{\rho}_{ABC} \Pi_{AC}) \\ &\leq P(F_B \hat{\rho}_{ABC} F_B^{\dagger}, \hat{\rho}_{ABC}) \leq \varepsilon', \end{aligned}$$

where we used the monotonicity of the purified distance (cf. Theorem 3.4) under projections. This concludes the proof. \square

A. VARIOUS LEMMAS

Appendix B

Properties of Quasi-Entropies

This appendix discusses properties of the quasi-entropies and relative Rényi Entropies introduced in Chapter 6.

B.1 Properties of Quasi-Entropies

The quasi entropies (cf. Def. 6.1) are well-defined in the sense that they are *covariant* under isometries on A and B .

Lemma B.1. *Let $U : \mathcal{H} \rightarrow \mathcal{H}'$ be an isometry. Then, for all f -quasi entropies and for all $A, B \in \mathcal{P}(\mathcal{H})$, we have*

$$S_f(A \parallel B) = S_f(UAU^\dagger \parallel UBU^\dagger).$$

Proof. Let $A = \sum_i \lambda_i |e_i\rangle\langle e_i|$ and $B = \sum_j \mu_j |f_j\rangle\langle f_j|$ with eigenvalues $\lambda_i \geq 0$, $\mu_j \geq 0$. Now, we write (see also [NO00])

$$S_f(A \parallel B) = \lim_{\xi \rightarrow 0} \sum_{i,j} (\mu_j + \xi) f\left(\frac{\lambda_i}{\mu_j + \xi}\right) |\langle e_i | f_j \rangle|^2.$$

The isometry U keeps eigenvalues and the scalar product $\langle e_i | f_j \rangle$ invariant. Furthermore, any zero eigenvalues introduced do not contribute since they lie in a space orthogonal to the image of U , where the summands vanish since $\lim_{\xi \rightarrow 0} \xi f(0) = 0$.¹ \square

Furthermore, it turns out that quasi-entropies have nice properties if the function f is chosen operator *concave*. In particular, the following

¹In Def. 6.1, we require that $f(0) \in \mathbb{R}$ is finite.

B. PROPERTIES OF QUASI-ENTROPIES

property holds.¹

Lemma B.2. *Let \mathcal{E} be a TP-CPM. Then, for all f quasi-entropies with f operator concave and for all $A, B \in \mathcal{P}(\mathcal{H})$, we have*

$$S_f(A \| B) \leq S_f(\mathcal{E}(A) \| \mathcal{E}(B)).$$

Proof. Any TP-CPM can be expressed as an isometry followed by a partial trace (cf. Lemma 2.4), hence, in conjunction with Lemma B.1, it remains to show the property for the partial trace operation. We will show this under the assumption that B is invertible and the result for general B will follow from the continuity (by definition) of S_f when $\xi \rightarrow 0$.

To show monotonicity under partial trace, we let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ with bases $\{|i\rangle_1\}$ and $\{|j\rangle_2\}$, respectively. We use the (unnormalized) fully entangled state $|\gamma\rangle = \sum_{i,j} (|i\rangle_1 \otimes |j\rangle_2) \otimes (|i\rangle_1 \otimes |j\rangle_2)$ in the product basis and its marginal $|\gamma\rangle_1 = \sum_i |i\rangle_1 \otimes |i\rangle_1$. It remains to show that $S_f(A \| B) \leq S_f(A_1 \| B_1)$, where $A_1 = \text{tr}_2(A)$ and $B_1 = \text{tr}_2(B)$.

Let us define a linear map $V : \mathcal{H}_1 \otimes \mathcal{H}'_1 \rightarrow \mathcal{H} \otimes \mathcal{H}'$ by

$$V := \sum_i \left(\sqrt{B} \left(\sqrt{B_1}^{-1} \otimes |i\rangle_2 \right) \right) \otimes \mathbb{1}_1 \otimes |i\rangle_2.$$

The map V is an isometry, i.e. $V^\dagger V = \mathbb{1}_{11}$ and satisfies

$$V(\sqrt{B_1} \otimes \mathbb{1}_1 |\gamma\rangle_1) = \sqrt{B} \otimes \mathbb{1}_{12} |\gamma\rangle. \quad (\text{B.1})$$

We have $\text{tr}_2(A^\text{T}) = A_1^\text{T}$, since the transpose is taken in the product basis. Hence, $V^\dagger(B^{-1} \otimes A^\text{T})V = B_1^{-1} \otimes A_1^\text{T}$. Next, we apply the Operator Jensen Inequality (Lemma 2.8) to get

$$V^\dagger f(B^{-1} \otimes A^\text{T})V \leq f(V^\dagger(B^{-1} \otimes A^\text{T})V) = f(B_1^{-1} \otimes A_1^\text{T}).$$

Finally, using (B.1), we recover $S_f(A \| B) \leq S_f(A_1 \| B_1)$ by taking the matrix element for $(\sqrt{B_1} \otimes \mathbb{1}_1) |\gamma\rangle_1$ on both sides of the inequality. \square

B.2 Properties of the Rényi Entropy

Here, we prove some general properties of the relative *Rényi entropies*, introduced in Chapter 6 as

$$S_\alpha(A \| B) = \frac{1}{1-\alpha} \log \text{tr} (A^\alpha B^{1-\alpha}).$$

¹This was essentially already shown in [Pet84] for the partial trace (see also [Hay06]) up to the continuity arguments above that allow us to define the quasi-entropies for non-invertible B . This extension (cf. Lemma B.1) is crucial, since even if B is invertible, $\mathcal{E}(B)$ is generally not if \mathcal{E} is an isometry. Hence, our contribution is to extend the monotonicity argument to general TP-CPMs.

B.2 Properties of the Rényi Entropy

Note that a similar quantity appears in quantum hypothesis testing [NO00, AMAV07] and in [Hay06, OP93, MD09], where alternative proofs of some of the following properties can be found.

Unlike their classical counterparts, the quantum relative (and conditional) min- and max-entropies cannot be recovered as special cases of α -entropies. However, it can be shown [KRS09] that

$$S_{1/2}(A\|B) = 2 \log \operatorname{tr}(\sqrt{A}\sqrt{B}) \leq 2 \log \operatorname{tr} |\sqrt{A}\sqrt{B}| = S_{\max}(A\|B). \quad (\text{B.2})$$

Furthermore, using the eigenvalue decompositions $A = \sum_i \lambda_i |e_i\rangle\langle e_i|$ and $B = \sum_j \mu_j |f_j\rangle\langle f_j|$, we have

$$S_{\infty}(A\|B) = \lim_{\xi \rightarrow 0} -\log \max_{\substack{i,j \\ \langle e_i | f_j \rangle \neq 0}} \frac{\lambda_i}{\mu_j + \xi} \leq S_{\min}(A\|B).$$

The entropies are additive, e.g. evaluation for an i.i.d. operator $A^{\otimes n}$ relative to another i.i.d. operator $B^{\otimes n}$ results in

$$S_{\alpha}(A^{\otimes n}\|B^{\otimes n}) = nS_{\alpha}(A\|B). \quad (\text{B.3})$$

The relative Rényi entropies decrease monotonically in α .

Lemma B.3. *Let $\alpha \geq \beta \geq 0$ and let $\rho \in \mathcal{S}_{=}(\mathcal{H})$, $\sigma \in \mathcal{P}(\mathcal{H})$ with $\operatorname{supp} \{\rho\} \subseteq \operatorname{supp} \{\sigma\}$. Then, $S_{\alpha}(\rho\|\sigma) \leq S_{\beta}(\rho\|\sigma)$.*

Proof. We first show that S_{α} decreases monotonically with increasing α by showing that its derivative is negative for all $\alpha \in [0, 1) \cup (1, \infty)$.

Using the (unnormalized) fully entangled state $|\gamma\rangle$, we define a purification $|\varphi\rangle := (\sqrt{\rho} \otimes \mathbb{1})|\gamma\rangle$ of A . Furthermore, we set $\bar{\alpha} = \alpha - 1$ and $X = (\rho \otimes \sigma^{-1})^{\top}$. It is easy to verify that, for $f : t \mapsto t \log t$,

$$\begin{aligned} S_{\alpha}(A\|B) &= -\frac{1}{\bar{\alpha}} \log \langle \varphi | X^{\bar{\alpha}} | \varphi \rangle \quad \text{and} \\ \frac{\partial}{\partial \alpha} S_{\alpha}(A\|B) &= \frac{1}{\bar{\alpha}^2} \log \langle \varphi | X^{\bar{\alpha}} | \varphi \rangle - \frac{1}{\bar{\alpha}} \frac{\langle \varphi | X^{\bar{\alpha}} \log X | \varphi \rangle}{\langle \varphi | X^{\bar{\alpha}} | \varphi \rangle} \\ &= \frac{f(\langle \varphi | X^{\bar{\alpha}} | \varphi \rangle) - \langle \varphi | f(X^{\bar{\alpha}}) | \varphi \rangle}{\bar{\alpha}^2 \langle \varphi | X^{\bar{\alpha}} | \varphi \rangle} \leq 0. \end{aligned}$$

The monotonicity follows from the convexity of f together with Jensen's inequality (cf. Lemma 2.7). \square

The Rényi entropies cannot decrease under simultaneous application of a CPM on both arguments. This is equivalent to a data-processing inequality for the conditional version of the entropies.¹

¹Conversely, it is easy to see that the data-processing inequality for the min- and max-entropy (cf. Theorem 5.7) imply monotonicity for S_{\max} and S_{\min} .

B. PROPERTIES OF QUASI-ENTROPIES

Lemma B.4 (Monotonicity of Rényi Entropy). *Let $\alpha \in [0, 2]$, let $A, B \in \mathcal{P}(\mathcal{H})$ and let \mathcal{E} be a TP-CPM. Then, $S_\alpha(A \| B) \leq S_\alpha(\mathcal{E}(A) \| \mathcal{E}(B))$.*

Proof. This follows directly from Lemma B.2 and the fact that $g_\alpha : t \mapsto t^\alpha$ is operator concave for $\alpha \in (0, 1)$ and operator convex for $\alpha \in (1, 2]$. (Also note that the pre-factor conveniently changes sign between these two domains.) In the limits $\alpha \rightarrow 0$ and $\alpha \rightarrow 1$, the property follows by continuity.¹ \square

We define conditional versions of these entropies as follows. For a state $\rho \in \mathcal{S}_\leq(\mathcal{H}_{AB})$, the conditional Rényi α -entropy of A given B is

$$H_\alpha(A|B)_\rho := S_\alpha(\rho_{AB} \| \mathbb{1} \otimes \rho_B).$$

This definition allows a *duality* relation for *pure* tri-partite states.

Lemma B.5. *Let $\rho \in \mathcal{S}_\leq(\mathcal{H}_{ABC})$ be pure and $\alpha \in [0, 1) \cup (1, 2]$. Then*

$$H_\alpha(A|B)_\rho + H_{2-\alpha}(A|C)_\rho = 0.$$

Proof. We write $\rho = |\vartheta\rangle\langle\vartheta|$ and note that the marginal states ρ_{AB} and ρ_C satisfy $\rho_{AB}|\vartheta\rangle = \rho_C|\vartheta\rangle$ and $\rho_B|\vartheta\rangle = \rho_{AC}|\vartheta\rangle$. Thus,

$$\begin{aligned} (1 - \alpha)H_\alpha(A|B)_\rho &= \log \text{tr}(\rho_{AB}^\alpha \rho_B^{1-\alpha}) = \log \langle \vartheta | \rho_{AB}^{\alpha-1} \rho_B^{1-\alpha} | \vartheta \rangle \\ &= \log \langle \vartheta | \rho_C^{\alpha-1} \rho_{AC}^{1-\alpha} | \vartheta \rangle = (\alpha - 1)H_{2-\alpha}(A|B)_\rho. \end{aligned}$$

The last equality follows from $\alpha - 1 = 1 - (2 - \alpha)$. \square

¹For the von Neumann entropy this property (and strong sub-additivity) also follows from the operator concavity of $h : t \mapsto -t \log t$, as noted in [NP05].

Index

- ε -ball, [45](#), [56](#), [71](#)
- ε -close, [53](#)
- ε -smooth, [45](#), [56](#)
- adjoint, [24](#)
- AEP, [17](#), [87](#), [127](#)
- algebra, *see* linear algebra
- angular distance, [49](#), [50](#)
- asymptotic equipartition, *see* AEP
- basis, [23](#)
 - orthonormal, [23](#)
- Bures metric, [49](#)
- chain rule, [17](#), [85](#)
- characterize, [16](#)
- coherent classical, [38](#), [40](#), [81](#), [154](#)
- concave, [42](#), [70](#), [159](#)
- conjugation, [31](#)
- contraction, [95](#)
- convergence parameter, [97](#)
- converse bound, [16](#), [101](#), [127](#), [130](#), [131](#), [133](#)
- convex, [42](#), [70](#)
- covariant, [159](#)
- CPM, [5](#), [31](#), [153](#)
- CQ, *see* state
- data processing, [11](#), [17](#), [70–72](#), [79](#), [127](#)
- decomposition, [23](#), [24](#)
 - eigenvalue, [29](#)
 - polar, [27](#), [53](#), [154](#)
 - Schmidt, [28](#), [63](#)
 - singular value, [27](#)
- dimension, [23](#), [34](#), [35](#)
- direct bound, [15](#), [100](#), [130](#), [133](#)
- direct sum, [23](#)
- distinguishing advantage, [47](#), [49](#), [132](#)
- duality, [17](#), [63](#), [72](#), [77](#), [162](#)
- effective overlap, *see* overlap
- eigenbasis, [29](#)
- eigenvalue, [29](#), [63](#)
- eigenvector, [29](#)
- embedding, [26](#)
- endomorphism, [25](#)
- entangled, [37](#)
- entanglement, [8](#), [9](#), [13](#), [14](#)
- entropy, [10](#), [45](#)
 - Hartley, [15](#), [16](#)
 - max-entropy, [15](#), [56](#), [60](#), [65](#)
 - min-entropy, [56](#), [58](#), [65](#), [155](#)
 - Rényi, [15](#), [55](#), [91](#), [95](#), [160](#)
 - Shannon, [10](#), [13](#), [93](#)
 - smooth, [15–17](#), [71](#), [73](#), [94](#), [103](#)
 - von Neumann, [12](#), [13](#), [55](#), [62](#), [88](#), [93](#), [103](#), [129](#)
- event, [10](#), [12](#)
- evolution, [39](#)
- extension, [38](#)
- fidelity, [47](#)
 - generalized, [48](#)
- generalized fidelity, *see* fidelity
- generalized inverse, [93](#), [94](#)
- guessing probability, [64](#), [84](#)
- Hamiltonian, [39](#)
- Hermitian, *see* operator
- hidden variable, [9](#)
- Hilbert space, [21](#), [27](#), [29](#), [35](#)
- homomorphism, [24](#)

INDEX

- i.i.d., 5, 11, 13, 87, 129
- i.i.d. limit, 12, 14, 17, 90, 127, 134
- identity, 25
- image, 25
- indicator, 23
- information theory, 9
- inner product, 21
 - Hilbert-Schmidt, 27
- inverse, 25
 - generalized, 25, 53, 154
- isometry, 26
 - partial, 26, 51
- isomorphism, 25, 27
 - Choi-Jamiołkowski, 32, 59
- kernel, 25
- lhs., 5
- linear algebra, 21
- marginal, 10, 37, 112, 153
- max-entropy, *see* entropy
- measurement, 12, 39
 - incompatible, 13, 103
 - projective, 40, 79
- measurement setup, 112
- metric, 22, 47, 50
- min-entropy, *see* entropy
- modulus, 27
- monotone, 42
- no-cloning, 9
- norm, 22
 - Schatten, 28
 - operator, 28
- normal, 29
- notation, 5, 6, 53
- observable, 103
- observer, 8, 11
 - ignorant, 8, 13, 18
 - omniscient, 8, 9, 18, 60
- one-shot, 14, 16, 17, 55–57, 88, 127, 134
- operator, 24
 - Hermitian, 29
 - identity, 25
 - Kraus, 33
 - linear, 24
 - positive, 30
 - positive semi-definite, 30
 - unitary, 26
- orthonormal, *see* basis
- overlap, 104, 108, 109
 - effective, 108, 112
- partial trace, 32
- partial transpose, 37
- POVM, 5, 40, 104
 - element, 40
- probability, 8, 10, 36, 40, 132
- projector, 25, 48, 51, 53
- purification, 38, 48, 61
- purified distance, 16, 46, 49, 72, 127
- QKD, 5, 127, 134
- quantum
 - cryptography, 9, 134
 - mechanics, 7, 34
- quantum channel, 39
- quantum state, 12, 13, 35
- qubit, 12, 35
- random variable, 10, 36, 40
- randomness extraction, 17, 127, 132
- rank, 25
- register, 12, 36, 39, 63, 72, 80, 104, 132
- representation
 - Kraus, 33, 153
- resource, 12, 14, 15
- rhs., 5
- SDP, 5, 41, 58, 59
- self-adjoint, 29
- semi-definite program, *see* SDP, 56
- separable, 37
- side information, 10, 11, 127, 132
- smoothing, 15, 56, 73, 74

smoothing parameter, 45
source compression, 14, 127
span, 22
state
 CQ, 5, 38, 64, 81, 104
 joint, 36, 38
 mixed, 36
 pure, 36, 38, 162
state merging, 14, 17
statistical distance, 16
Stinespring dilation, 33, 40, 80, 85
sub-multiplicative, 29
sub-unital, 32
super-operator, 31
support, 25
surprisal, 10, 11, 13, 16
symmetric, 48
system, 12, 35
 joint, 36

tensor product, 23, 30
TP-CPM, 5, 31, 39, 51, 59, 72, 79,
 81, 153
trace, 26, 32
 non-increasing, 32
 preserving, 31
trace distance, 46, 47
trace non-increasing, 51
transpose, 25
typical set, 87, 89

UCR, *see* uncertainty
Uhlmann's theorem, 48, 52, 61
uncertainty, 10
 principle, 8
 relation, 5, 103
uncertainty relation, 13
unital, 32, 59, 79
unitarily invariant, 29, 76, 154
unitary, 39

vector space, 21
von Neumann entropy, *see* entropy